


```

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.;
; Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69257
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-196 <KLE>
A;Cross-references: UNIPROT:O30175; GB:AE001102; GB:AE000782; NID:g2689425; PIDN:AAB9116
C;Superfamily: dTMP kinase

Query Match      6.6%; Score 155; DB 2; Length 196;
Best Local Similarity 28.6%; Pred. No. 0.00018;
Matches 61; Conservative 37; Mismatches 75; Indels 40; Gaps 10;

QY 254 VVAIEGLDAGTKTTVTQSVADSL-----KAVLLKSP-PSICIGQWRKIFDDEPTTIIRAFY 307
Db 2 LIAVEGDGAGKTTIAIAVIELLKEGYKVKVLPBGDSKFG--KKIKSSEERISPEEL 59

QY 308 SLGNYIVASET-----AKSPVIVDRYHSTATYATATEVSGGLQHLPPAHPVY 360
Db 60 EL--FLKDREIDARENILPALQSGVAVVMDRYFYSNIAQSGIDARL-----IR 108

QY 361 QWPEDLL-KPDILLITVSPERLQROGMKTRAEAELEANSVFRQKVMYSQRMEN 419
Db 109 EMNEKIAPKPDILTILDEPITALERVKRGLSPFEKLD-----YLRKVRKGFLENAD 162

QY 420 PGCHVVDASPSRE-----KVLQTVLSLIONS 445
Db 163 ETTVVVDASKPLEVKEVRKVIESFLNKKNS 195

RESULT 3
F82788
thymidylate kinase XF0580 [imported] - Xylella fastidiosa (strain 945c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: F82788
R;anonymus: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82788
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <SIM>
A;Cross-references: GB:AE003904; GB:AE003849; NID:g9105433; PIDN:AAP83390.1; GSPDB:GN001
A;Experimental source: strain 945c
R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Cartaro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigz
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr. V.R.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
A;Genetics:
A;Gene: XF0580
C;Superfamily: dTMP kinase

Query Match      6.5%; Score 154.5; DB 2; Length 217;
Best Local Similarity 27.7%; Pred. No. 0.00022;
Matches 59; Conservative 32; Mismatches 71; Indels 51; Gaps 9;

```

```

QY 254 VVAIEGLDAGTKTTVTQSVADSLKAVLL-----KSPSPSCIGQWRKIF-----DD 297
Db 11 LVAIEGDGAGKTTTALSKLURKUGVLETVVSKPTN--GPGWMLRQSAATGRFSPDEE 68

QY 298 EPTTIIRAFYSLGNYIVASEIAKESAKSPVIVDRYHSTATYATATEVSGGLQHLPPAHH 357
Db 69 EVDVLLDRQRHVEDLIVPMIGRAV--VILDRYFSPMVAY-----QGAAGL----- 113

QY 358 PVYQWPEDELL-----KPDILLITVSPERLQROGMKTRAEAELEANSVFRQK 409
Db 114 -----PVDALLEANAFAPRPDVLILLDPPAIGLQRIWERG--STPNHFETTEN---LSR 163

QY 410 VMSYQRMENPGCHVVDASPSREKVLQTVLSLI 442
Db 164 CRDIFLALPLPSKRVKVIDATANAETVLSAALALV 196

RESULT 4
T41553
thymidylate kinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41553
R;Wood, V.; Rajandream, M.A.; Bartrell, B.G.; Seeger, K.; Harris, D.
submitted to the EMBL Data Library, June 1998
A;Reference number: 222001
A;Accession: T41553
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-210 <WOO>
A;Cross-references: UNIPROT:P36590; EMBL:AL023794; PIDN:CAA19357.1; GSPDB:GN00068; SPDB:
A;Experimental source: strain 972h-; cosmid c70
C;Genetics:
A;Gene: SPDB:SPCC70.07c
A;Map position: 3
A;Introns: 45/1; 110/3
C;Superfamily: dTMP kinase

Query Match      6.1%; Score 144.5; DB 2; Length 210;
Best Local Similarity 26.1%; Pred. No. 0.0012;
Matches 58; Conservative 39; Mismatches 76; Indels 49; Gaps 11;

QY 248 QKGFOVVAIEGLDAGTKTTVTQSVADSL-----KAVLLKSPSPSCIGQWRKIFD----- 296
Db 3 KQNGRIIVIEGLDRSGKSTOCQLLVKLSQHEKAELEKFPDRTTAIGKKIDDYLKESV 62

QY 297 ---DEPTTIIRAFYSLGNYIVASIAKESAKS-PVIVDRYHSTATYATATEVSGGLQHL 352
Db 63 QLNDQ---VIHLFSANRWETIQIYEQINKGVTCILDY----AFSGIAFSAAGLD-- 113

QY 353 PPAIHFPVQW-----PEDLLKPDILLITVSPERLQROGMKTRAEAELEANSVFRQ 408
Db 114 -----WENCKSPDRGLPRDLVIFLNVDP--RIAAATRGQYGBERVEKIEMQ-----E 158

QY 409 KVMSYQRMEN---NPGCHVV--DASPSREKVLQTVLSLION 444
Db 159 KVLKNFORLQKFEFREGLEFILDASSSLEVDHVSQIVDLVSN 200

RESULT 5
thymidylate kinase (tmk-1) [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: F90227
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensesen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90227

```


QY 302 I-RRAFYSLGNYIVASEIAKESAK-----SPVIVDRYWHSTATYAIATEVSGGLQH 351

Db 60 DGSKISYBAEALLFAADRAEHVKKIIILPALSEKGVVICDRFYSSL----- 105

QY 352 LPPAHPHYQWPEDL-----LKPDLILLITVSPERLQRLQGRGMKTRERAE 399

Db 106 -----AYQWARGLDLNLWLVQVNSFARPDLAILLDLFVKESURRIKRG-----TITE 153

QY 400 LEANSVFRQKVMYSORMEN--PGCHVVVDASPSREKVLQTVLSLION 444

Db 154 FDKIVELQKRVHNYIKLAEMFPMRIUNALSSIEDIHSDIVALKH 200

RESULT 9

A82128

thymidylate kinase VC2016 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: A82128

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82128

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-212 <HEI>

A:Cross-references: UNIPROT:Q9KQ12; GB:AE004276; GB:AE003852; NID:g9656555; PIDN:AAF9516

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2016

A:Map position: 1

C:Superfamily: dTMP kinase

Query Match 5.3%; Score 124.5; DB 2; Length 212;

Best Local Similarity 24.1%; Pred. No. 0.04;

Matches 53; Conservative 36; Mismatches 88; Indels 43; Gaps 7;

QY 255 VAIEGLDAGTKTTVQSV-----ADSLKAVLLKSPSCIGQW 291

Db 6 IVIEGLEAGKSTAQVVVETLQONGIDHITRTREPGGTLAEKLRALVKEHPG----- 60

QY 292 RKIFD-DEPTIIRAFYSLGNYIVASEIAKESAKSPVIVDRYWHSTATYAIATEVSGGLQ 350

Db 61 EELQDITELLIVYARVOLVENVIKPALARGEW---VVGDRHDSWSSQAYQ-----GGGRQ 112

QY 351 HLPFAHPVYQWPEDLKPDILLITVSPERLQRLQGRGMKTRERAELEANSVFRQKV 410

Db 113 IAPSTMQSLKQTALGDFKPDITLYLIDIPKLGLEARGGLDRIEKWDIS-----FFERA 168

QY 411 EMSYORMENPGCHVV--DASPSREKVLQTVLSLIONSFSE 448

Db 169 RERYLELANSDDSVVMDAAQSIQVTAIDIRRALQDWLSQ 208

RESULT 10

S28955

dTMP kinase (EC 2.7.4.9) - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004

C:Accession: S28955

R:Abigail, L.T.; Yeh, Y.I.S.; Jong, A.Y.

Biochim. Biophys. Acta 1132, 222-224, 1992

A:Title: Functional and structural conservation of Schizosaccharomyces pombe dTMP kinase

A:Reference number: S28955; MUID:93003330; PMID:1327149

A:Accession: S28955

A:Molecule type: mRNA

A:Residues: 1-210 <AEA>

A:Cross-references: UNIPROT:P36590; EMBL:X65868; NID:g4973; PIDN:CAA46698.1; PID:g4974

C:Superfamily: dTMP kinase

C:Keywords: ATP; nucleotide binding; nucleotide biosynthesis; P-loop; phosphotransferases

F:14-21/Region: nucleotide-binding motif A (P-loop)

Query Match 5.2%; Score 123.5; DB 2; Length 210;

Best Local Similarity 24.3%; Pred. No. 0.047;

Matches 57; Conservative 36; Mismatches 67; Indels 75; Gaps 12;

QY 248 QKGFQVVAIEGLDAGTKTTVQSVADSL-----KAVLLKSPSCIGQWRIKIFDD----- 297

Db 3 KQNGRLIVIEGLDRSGKSTOCQLLVKLIILNMKRLKLFKPDRTTAIGKKI-DDYLLES 61

QY 298 -----EPTIIRAFYSLGNYIVASEIAKESAKSPVIVDRYWHSTATYAI 340

Db 62 VQLNDQVTHLLFSANRWEPSIYYRA-----NQQRNCILDY-----AFSG 102

QY 341 IATEVSGGLQHLPPAHPHYQW---PEDLLKPDILLITVSPERLQRLQGRGMKTR 396

Db 103 IAFSAAKGLD-----WEWCKSPDRGLTRPDLVIFLNVDP--RIAATRQYGEERYE 151

QY 397 EAELEANSVFRQKVMYSORME---NPGCHVV--DASP-SREKVLQTVLSLION 444

Db 152 KIEWQ-----EKVLKNLQRLQKFRBEGLEFILLDASSYADEVDVDSQIVDLVSN 200

RESULT 11

D90518

thymidylate kinase (dTMP kinase) [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: D90518

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: D90518

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <KUR>

A:Cross-references: UNIPROT:Q98RF7; GB:AL445566; PID:g14089465; PIDN:CAC13225.1; GSPDB:

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 0520

A:Genetic code: SGC3

C:Superfamily: dTMP kinase

Query Match 5.2%; Score 123; DB 2; Length 234;

Best Local Similarity 25.5%; Pred. No. 0.059;

Matches 56; Conservative 43; Mismatches 75; Indels 46; Gaps 14;

QY 255 VAIEGLDAGTKTTVQSVADSLKAVLLKS--PPSCIGQW-----RK-----IFDDE 298

Db 16 ITFEGIDASGT-----SLLAKLKAHVQKNVQKNLGKCTFTWEPGGRKSPELQIRHLINKE 71

QY 299 PTI--IRRAF-YSLG-----NYIVASEIAKESAKSPVIVDRYWHSTATY-AIATEVSGGL 349

Db 72 SNLSPIAEAFUYSSARRIHLDDKVLPLNKNKV--VPCDRFVDSVPAYQAFGRDL--GF 126

QY 350 QHLPFAHPHYQWPEDLKPDILLITVSPERLQRLQGRGMKTRERAELEANSVFRQK 409

Db 127 EKIKLN---ELATDKKYPDITFIKISVEESMERKAR--QDEDRLEKEENS-FYQK 179

QY 410 VEMSYORMEN-----PGCHVVVDASPSREKVLQTVLSLION 444

Db 180 VIKGYDFLASYPKQKIFVIDASKNQBEIFESVLKILKD 219

RESULT 12

S66058

thymidylate kinase tmk - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S66058; D69724

R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.

DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chrom
A;Reference number: S65967; MUID:96051385; PMID:7584024
A;Accession: S66058
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-212 <OGA>
A;Cross-references: UNIPROT:P37537; EMBL:D26185; NID:9467326; PIDN:BAA05264.1; PID:g4674
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
R;Kunst, F.; Ogasawara, S.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brucher, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
A.; Koether, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A.; Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A.; Yoshikawa, H.F.; Zunstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D69724
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-212 <KUN>
A;Cross-references: GB:299104; GB:AL009126; NID:g2632267; PIDN:CAB11804.1; PID:g2632295
A;Experimental source: strain 168
C;Genetics: tmk
C;Superfamily: dTWP kinase
C;Keywords: nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)

Query Match 5.1%; Score 120; DB 2; Length 212;
Best Local Similarity 25.0%; Pred. No. 0.087;
Matches 53; Conservative 36; Mismatches 97; Indels 26; Gaps 8;
Qy 255 VAIEGLDAGTKTTVTSQVADSLKA-----VLLKSPSCIG---QWRKIFDDEPTII-----302
Db 6 ITFEGEGAGTKTTLQEIKNILTAEGLVQVATREPGGHDIAEQREVILNENNILMDPKT 65
Qy 303 RRAYSIGNIVASEIAKESAKSPVIV--DRYWHSTATYATATEVSGGLQHLPPAHHPVY 360
Db 66 EALLIYAARROHLVEKVKPALEQGFVLCDFIDSPAYQGYARGLGIDEVL-----SIN 120
Qy 361 QWPEDLLKPDILILLTVSPERLQGRGMEKREAELEANSV-PRQKVMYSYQRMEN 419
Db 121 EFAIGDMPHVTYVFSIDPEGLKRIYANG---SREKNRLDLEKIDFHTKVQEGYQELMK 177
Qy 420 ---PGCHVVDASPREKVLQTVLSIQNSFSSE 448
Db 178 RFPERFHSVDAGSKQLVQDVLKVIDEALKK 209

RESULT 13
KIV25W
dTMP kinase (EC 2.7.4.9) - vaccinia virus
N;Alternate names: A48R protein; SalFlir protein; thymidylate kinase
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: E42522; JQ1786; S06181; T37434
R;Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A;Title: Appendix to "The complete DNA sequence of vaccinia virus".
A;Reference number: A42501
A;Accession: E42522
A;Molecule type: DNA
A;Residues: 1-204 <GOE>

A;Cross-references: UNIPROT:P13410; GB:M35027; NID:g335317; PIDN:AAA48180.1; PID:g335526
A;Experimental source: strain Copenhagen
R;Smith, G.L.; Chan, Y.S.; Howard, S.T.
J. Gen. Virol. 72, 1349-1376, 1991
A;Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right i
A;Reference number: JQ1767; MUID:91259063; PMID:2045793
A;Accession: JQ1786
A;Molecule type: DNA
A;Residues: 1-204 <SMI>
A;Cross-references: DDBJ:D11079; NID:g222717; PIDN:BAA01822.1; PID:g222737
A;Experimental source: strain strain WR
R;Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A;Title: The complete DNA sequence of vaccinia virus.
A;Reference number: A42531; MUID:91021027; PMID:2219722
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
R;Smith, G.L.; de Carlos, A.; Chan, Y.S.
Nucleic Acids Res. 17, 7581-7590, 1989
A;Title: Vaccinia virus encodes a thymidylate kinase gene: sequence and transcriptional
A;Reference number: S06181; MUID:90016845; PMID:2552411
A;Accession: S06181
A;Molecule type: DNA
A;Residues: 1-204 <SM2>
A;Cross-references: EMBL:X16259; NID:g62231; PIDN:CAA34345.1; PID:g622232
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.
submitted to the EMBL Data Library, March 1997
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) stra.
A;Reference number: Z20877
A;Accession: T37434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-204 <ANT>
A;Cross-references: EMBL:U94848; PIDN:AAB96539.1
A;Experimental source: strain Ankara
C;Genetics: tmk
A;Note: MVA161R
C;Superfamily: dTWP kinase
C;Keywords: ATP; nucleotide binding; nucleotide biosynthesis; P-loop; phosphotransferas
F;11-18/Region: nucleotide-binding motif A (P-loop)
F;17/Binding site: ATP (Lys) #status predicted

Query Match 5.1%; Score 119.5; DB 1; Length 204;
Best Local Similarity 28.0%; Pred. No. 0.09;
Matches 51; Conservative 28; Mismatches 66; Indels 37; Gaps 8;
Qy 254 VVAIEGLDAGTKTTVTSQVADSLKAVLLK--SPSPCIGQWRKIFDDEPTIIRAFYSLGN 311
Db 6 LIVFEGLDKSGKTTQCNWIMESIPANTIKYLPFRQSVITGKMIDDYLT--RKTYN--D 61
Qy 312 YIV-----ASEIAKESAKS-PVIVDRYWHSTATYATATEVSGGLQHLPPAHHPV 359
Db 62 HIVNLLFCANRWFASFQIEQEGITLIVDRYAFSGVAYAAAGASMTLSK-----113
Qy 360 QWPEDLLKPDILILLTVSPERLQGRGMEKREAELEANSVPRQKVMYSYQRMEN 419
Db 114 -SYSGLPKPDLVIFLE--SGSKENRNNGE-----EIVDVTFOOKVLEFKKMLE 162
Qy 420 PG 421
Db 163 EG 164

RESULT 14
P64336
dTMP kinase (EC 2.7.4.9) - Methanococcus jannaschii
N;Alternate names: thymidylate kinase
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: F64336
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, A.; Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* A; Reference number: A64300; PMID:96337999; PMID:8688087

A;Accession: F64336
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-188 <BUL>

A;Cross-references: UNIPROT:Q57741; GB:U67484; GB:L77117; NID:q1591009; PIDN:AAB98278.1;

C;Genetics:

A:Map posit:
A:start code

C;Superfamily: dTMP kinase
C;Keywords: nucleotide binding site

C;Keywords: nucleotide binding site
F;11-18/Region: nucleotide-binding site

Query Match	Score	DB 1	DB 2	Length
5.0%	Score 118.5	DB 2	Length	

Best Local Similarity 23.9%; Pred. No. 0.096;
Matches 47; Conservative 41; Mismatches 8

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 20:21:58 ; Search time 157 Seconds
(without alignments)
1645.498 Million cell updates/sec

Title: US-10-681-223-2
Perfect score: 2362
Sequence: 1 MAFARLLRGLSGPLIGRR.....SREKVLQTVLSLQNSFSEP 449

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1684	71.3	417	2 Q6PFG7	Q6pfg7 mus musculus
2	1684	71.3	417	2 AAH57565	AAH57565 mus muscu
3	1552	65.7	395	2 Q9DC34	Q9dc34 mus musculus
4	1466	62.1	303	2 Q6ZRU2	Q6zru2 homo sapien
5	1466	62.1	303	2 BAC87217	Bac87217 homo sapi
6	677.5	28.7	431	2 Q62316	Q62316 mus musculu
7	301	12.7	274	2 Q60970	Q60970 leishmania
8	291	12.3	58	2 Q96AU8	Q96au8 homo sapien
9	169.5	7.2	227	1 KTHY_XANAC	K96af7 xanthomonas
10	166.5	7.0	227	1 KTHY_XANCP	Q9p3y6 xanthomonas
11	155	6.6	196	1 KTHY_ARCFU	Q30175 archaeoglob
12	154.5	6.5	208	1 KTHY_XYLFA	Q9pfs7 xylella fas
13	148.5	6.3	208	1 KTHY_XYLFT	Q87b89 xylella fas
14	145	6.1	202	2 Q6L016	Q6l016 picophilius
15	144.5	6.1	210	1 KTHY_SCHPO	P36590 schizosacch
16	143	6.1	189	1 KTH1_SULSO	Q9uxg7 sulfolobus
17	142	6.0	204	2 Q72NL9	Q72nl9 leptospira
18	142	6.0	204	2 AAS71368	Aas71368 leptospir
19	140	5.9	209	1 KTHY_PASWU	Q9cke9 pasteurella
20	137.5	5.8	263	2 Q9LEB4	Q9leb4 arabidopsis
21	135.5	5.7	210	1 KTHY_PYRFU	Q9u071 pyrococcus
22	134	5.7	191	2 Q8SS22	Q8ss22 encephalito
23	131.5	5.6	205	2 Q6N1S4	Q6nis4 corynebacte
24	131.5	5.6	205	2 CAE49210	Cae49210 corynebac
25	130.5	5.5	199	1 KTHY_TROW8	Q93hb8 tropheryma
26	130.5	5.5	199	1 KTHY_TROWA	Q93fm3 tropheryma
27	130.5	5.5	206	1 KTHY_TMETW	Q9pxv5 methanosarc
28	129.5	5.5	211	2 Q6F0B1	Q6foel mesoplasma
29	129	5.5	211	2 Q9AOD5	Q9aod5 lactococcus
30	129	5.5	213	2 Q6MU15	Q6mu15 mycoplasma
31	129	5.5	213	2 CAE76699	Cae76699 mycoplasma

32 127.5 5.4 130 1 KTH1_SULTO Q975e6 sulfolobus

33 127.5 5.4 205 1 KTHY_PYRAB Q9vie9 pyrococcus

34 127 5.4 204 2 Q6CC22 Q6cc22 yarrowia li

35 127 5.4 1082 1 IPO4 MOUSE Q8vi75 mus musculus

36 126 5.3 210 1 KTHY_HAEIN P44719 haemophilus

37 126 5.3 218 2 Q6FSB3 Q6fsb3 candida gla

38 125.5 5.3 205 1 KTHY_PYRHO O59366 pyrococcus

39 124.5 5.3 212 1 KTHY_VIBCH Q9xq12 vibrio chol

40 124 5.2 208 2 Q6CW62 Q6cw62 kluyveromyc

41 124 5.2 227 2 Q8OHT9 Q8oht9 vaccinia vi

42 123.5 5.2 213 1 KTHY_ECOL6 Q8fin9 escherichia

43 123 5.2 227 2 Q8ODS7 Q8ods7 cowpox viru

44 123 5.2 234 1 KTHY_MYCFU Q98rf7 mycoplasma

45 123 5.2 1017 1 UB35_HUMAN Q9p2h5 homo sapien

ALIGNMENTS

RESULT 1

Q6PFG7 PRELIMINARY; PRT; 417 AA.

AC Q6PFG7; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Tyki protein.

GN Name=Tyki;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6; TISSUE=Mouse;

RX MEDLINE=22398257; Pubmed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalek U., Smalitz D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6; TISSUE=Mouse;

RC Strausberg R.;

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC057565; AAH57565.1; -

DR InterPro; IPR000662; Thymidylate kin.

DR Pfam; PF02223; Thymidylate kin; 1.

SQ SEQUENCE 417 AA; 46646 MW; 978DE9D0AA83F12A CRC64;

Query Match 71.3%; Score 1684; DB 2; Length 417;

Best Local Similarity 77.1%; Pred. No. 3.5e-115;

Matches 323; Conservative 38; Mismatches 54; Indels 4; Gaps 2;

Qy 27 MAPPCRVLPLEPDCITLAHFALGADAPGDADAPPRLAALLGPPERSYSLCVPTPDAGCG 86

Db 1 MARPRRTVELPDCSLTHFVLG-DA---TTHRDLARLAELGPPGRSVALCVLPAPGEGCG 56

```

QY 87 ARVAAARLHORLLHQLRRGPPRCQLRLRLCYCGGAGGAGGQOQFLRRDPLDDPTQAL 146
DB 57 PRVQAARVHRLQLRRGPPRCQLRLRLCYCGGAGGAGGQOQFLRRDPLDDPTQAL 116
QY 147 LELGACQEAAPRHGLGFEADPRQLMORLWEVGDGRLOVGCACVVPVPPPLHPVVPD 206
DB 117 LQLLGSQEAARPOLAFQADSQGLLWOLWELQDGRQVQVDCACVLPQAQPHLPLFPD 176
QY 207 LPSSWFPDREAARAVLEECTSFIPPEARAVLDLVDQCPKQIQKGFQVVAIEGLDATGKT 266
DB 177 LLNSAVFQDRDAARAVLEECTSFIPPEARAVLDLVDQCPKQIQKGFQVVAIEGLDATGKT 236
QY 267 TVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFYSLGNIVASIAKESAKSP 326
DB 237 TLTSQSVESLKAVLLQSPPCISQWRKIFDDEPTIIRAFYSLGNIVASIAKESAKSP 296
QY 327 VIVDRYWHSTAYATATVSGGLQHLPPAHPVQWQPEDLLKPDILLLTSPERLQRL 386
DB 297 VIVDRYWHSTAYATATVSGGLQHLPPAHPVQWQPEDLLKPDILLLTSPERLQRL 356
QY 387 QGRGMEKTREAELEANSVFRQKVMYSORMENPCHVVDASPSREKVLQTVLSLIONS 445
DB 357 QGRGMEKTREAELEANSVFRQKVMYSORMENPCHVVDASPSREKVLQTVLSLIONS 415

RESULT 2
AAH57565 PRELIMINARY; PRT; 417 AA.
ID AAH57565
AC AAH57565;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Tyki protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Mouse;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Mouse;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC057565; AAH57565.1; -
SQ SEQUENCE 417 AA; 46646 MW; 978DE9D0AA83F12A CRC64;

```

Query Match 71.3%; Score 1684; DB 2; Length 417;
 Best Local Similarity 77.1%; Pred. No. 3.5e-115;
 Matches 323; Conservative 38; Mismatches 54; Indels 4; Gaps 2;

```

QY 27 MAPPCRFVLELPCDTLAHFAFGADAPGDADAPDRLAALLGPPERSYSLCVPTFDAGCG 86
DB 1 MARPRFTVELPDCSLTHFVLG-DA---TDRDARLAELLGPPGRSYALCVPLAPGEGCG 56
QY 87 ARVAAARLHORLLHQLRRGPPRCQLRLRLCYCGGAGGAGGQOQFLRRDPLDDPTQAL 146
DB 57 PRVQAARVHRLQLRRGPPRCQLRLRLCYCGGAGGAGGQOQFLRRDPLDDPTQAL 116
QY 147 LELGACQEAAPRHGLGFEADPRQLMORLWEVGDGRLOVGCACVVPVPPPLHPVVPD 206
DB 117 LQLLGSQEAARPOLAFQADSQGLLWOLWELQDGRQVQVDCACVLPQAQPHLPLFPD 176
QY 207 LPSSWFPDREAARAVLEECTSFIPPEARAVLDLVDQCPKQIQKGFQVVAIEGLDATGKT 266
DB 177 LLNSAVFQDRDAARAVLEECTSFIPPEARAVLDLVDQCPKQIQKGFQVVAIEGLDATGKT 236
QY 267 TVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFYSLGNIVASIAKESAKSP 326
DB 237 TLTSQSVESLKAVLLQSPPCISQWRKIFDDEPTIIRAFYSLGNIVASIAKESAKSP 296
QY 327 VIVDRYWHSTAYATATVSGGLQHLPPAHPVQWQPEDLLKPDILLLTSPERLQRL 386
DB 297 VIVDRYWHSTAYATATVSGGLQHLPPAHPVQWQPEDLLKPDILLLTSPERLQRL 356
QY 387 QGRGMEKTREAELEANSVFRQKVMYSORMENPCHVVDASPSREKVLQTVLSLIONS 445
DB 357 QGRGMEKTREAELEANSVFRQKVMYSORMENPCHVVDASPSREKVLQTVLSLIONS 415

RESULT 3
Q9DC34 PRELIMINARY; PRT; 395 AA.
ID Q9DC34
AC Q9DC34;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,
DE clone:120004E04 product:thymidylate kinase family LPS-inducible
DE member, full insert sequence.
GN Name=lyxi;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

```

RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Katsunai T., Tashiro H., Itoh M.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishigaki T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanaoka T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami T., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK004595; BAB23396.1; -
DR HGSP; P00572; 1TMK.
DR MGD; MGI-99830; Tyki.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004798; F:thymidylate kinase activity; IEA.
DR GO; GO:0006233; P:dTTP biosynthesis; IEA.
DR GO; GO:0006235; P:dTTP biosynthesis; IEA.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; I.
KW Kinase.
SQ SEQUENCE 395 AA; 44085 MW; 62F338B649983DF6 CRC64;
Query Match 65.7%; Score 1552; DB 2; Length 395;
Best Local Similarity 77.8%; Pred. No. 1.6e-105;
Matches 294; Conservative 36; Mismatches 48; Indels 0; Gaps 0;
Qy 68 PPSRSLCVPTDPDAGCGARVRAARLHQLRRLGPGQAGGA 127
Db 16 PQGAATRLCVLAPGEGCGPRVQARVHRLQLLRGFLQRCQLKLLGPGDQAGEA 75
Qy 128 QQGLRLDPLDDPTROALLLELGGACQAPRPHLGEFEADPRGQLWQLRWEVQGRRLQV 187
Db 76 QHGLRLDPCDHPDTRRLDQLLGGSCQEAAPQJAEFOADSGQLLWQLRWEVQGRRLQV 135
Qy 188 GCAQVVPVPEPLHPVPPDPSVVPDREARAVLEECTSFIPPEARAVLDLVDQCPKQI 247
Db 136 DCACVLPQAEPLHPPLDPLNSAVFQDRAARAVLEECTSFIPPEARAVLDLVDQCPKEV 195
Qy 248 QKGFQVVAIEGLDAGTKTTVQSVADSLKAVLLKPPSCIGQWRKIFDDEPTTIIRAFY 307
Db 196 QKGFQVVAIEGLDAGTKTTVQSVADSLKAVLLKPPSCIGQWRKIFDDEPTTIIRAFY 255
Qy 308 SLGNVIVASEAKESAPVTVDRVWHSTATVAIATEVSGGLQHPRAHHPVQWPEDLL 367
Db 256 SLGNVIVASEAKESAPVTVDRVWHSTATVAIATEVSGGLQHPRAHHPVQWPEDLL 315
Qy 368 KPDILLITVSPERLQLRGMEKTRAELEANSVFRQKVENSMENPGCHVYDA 427
Db 316 KPDVLLITVNSEERVRLQGRQEKTEAELEANNVFRQKVENSMENPNSCHLYDA 375

Qy 428 SPSREKVLQTVLSLIONS 445
Db 376 SPSRETVLQKVLLEIQSS 393
RESULT 4
Q6ZRU2 PRELIMINARY; PRT; 303 AA.
ID Q6ZRU2
AC Q6ZRU2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46098.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.,
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK127983; BAC87217.1; -
DR GO; GO:0016301; F:kinase activity; IEA.
KW Kinase.
SQ SEQUENCE 303 AA; 32645 MW; 4C7D0715B7B970C6 CRC64;
Query Match 62.1%; Score 1466; DB 2; Length 303;
Best Local Similarity 99.6%; Pred. No. 2.3e-99;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAFARRLLRGLPLGRGVCAAGVAPCRFVLELPCDTLAHFALGADAPGDADPDP 60
Db 1 MAFARRLLRGLPLGRGVCAAGVAPCRFVLELPCDTLAHFALGADAPGDADPDP 60
Qy 61 RLAAALLGPPERSYSLCVPTDPDAGCGARVRAARLHQLRRLGPGQAGGA 120
Db 61 RLAAALLGPPERSYSLCVPTDPDAGCGARVRAARLHQLRRLGPGQAGGA 120
Qy 121 GGQAGGACQGFLLRDLDDPTROALLLELGGACQAPRPHLGEFEADPRGQLWQLRWEVQ 180
Db 121 GGQAGGACQGFLLRDLDDPTROALLLELGGACQAPRPHLGEFEADPRGQLWQLRWEVQ 180
Qy 181 DGRRLQVCAQVVPVPPPLHPVPPDPSVVPDREARAVLEECTSFIPPEARAVLDL 240
Db 181 DGRRLQVCAQVVPVPPPLHPVPPDPSVVPDREARAVLEECTSFIPPEARAVLDL 240
Qy 241 DQCEKQIKGKGFQVVAIEGLDAGTKTTVQSVADSLKAV 279
Db 241 DQCEKQIKGKGFQVVAIEGLDAGTKTTVQSVADSLKAV 279
RESULT 5
BAC87217 PRELIMINARY; PRT; 303 AA.
ID BAC87217
AC BAC87217
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FUJ46098 fls. clone TESTI2021654, weakly similar to Mus musculus
DE thymidylate kinase family LPS-inducible member (Tyki).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

	Query Match	28.7%	Score 677.5;	DB 2;	Length 431;
	Best Local Similarity	58.9%;	Pred. No. 2.7e-41;		
	Matches 142;	Conservative 24;	Mismatches 45;	Indels 33;	Gaps 5
Qy	225 ECTSFPEARVLVDVDCPQIOKGKGVVAIEGLDGTGKTWTVSADSLKAVLLKSP	284	:	:	:
Dd	157 QCTSFPEARVLVDVDCPKEVGKGFVIAEGLDGTGKITTLQHFKPSLSRSLSSYGRH	216	:	:	:
Qy	285 PSCIGQRWKIPDD-----EPTIIRAPFYSLGNIIVASEIAKESAKSPVIIDRVYWH	334	:	:	:

Query Match	12.7%;	Score 301;	DB 2;	Length 274;
Best Local Similarity	32.0%;	Pred. No. 6.6e-14;		
Matches	85;	Conservative 39;	Mismatches 108;	Indels 34;
				Gaps 7
Qy	184	RLOVGCQVVPVPB---	PPLHPVDPDLPSSVFPDREARAVLEBCTSFIPCARAVLDIV	240
Db	19	RSRLDCLKVLHAEQSSFPFPWTPD-----	CSNFVRQAEALSTSA	59
Qy	241	DQCKPQIQKGKQFQVAIEGLD	ATGKTTVTSQVADSLKAVLLKSPSC	TGQWRKIFDDEPT 300
Db	60	SATAAR----	KNPVIWVEGLDGTGKTLVTRTAEKLSGVAISTPPQPTERTNTF	RGQBE 115
Qy	301	IIRAFYSLGNYYIVASEIAKESAP	VPVDVRYWYHSTATVATATEVSGGI-QHLPAHPV	359
Db	116	AVARAFYSANYIAEGLLAASQ	SVVVVDKRWGCTCMAIA--	NGCLYSLPESGRAV 172
Qy	360	QYWPEDLLKPDLLILLTVS	PERIQIQGRGMEKTRAEAELEANSVFQKVMYS	VQRMEN 419
Db	173	YRWPEDLPAFDAGFLLCV	DEAVRVARIRRAPEDA-ERRR	SSQREMRQVAMEAYRR--- 228
Qy	420	PGCHVDASPSREKVLQTVLS	LIQNS	445
Db	229	TNMLIEVAPSYRVAVNSIL	RLPES	254
RESULT	8			

```

Q96AL8
ID Q96AL8 PRELIMINARY; PRT; 58 AA.
AC Q96AL8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RW EMBL; BC016969; AAH16969.1;
FT NON_TER
SQ SEQUENCE 58 AA; 653 MW; C12029FED37F4BAC CRC64;

Query Match 12.3%; Score 291; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EKTREAELEANSVFQKVMYSQRMENPCGHVVDASPSREKVLQTVLSLQNSFSSEP 449
Db 1 EKTREAELEANSVFQKVMYSQRMENPCGHVVDASPSREKVLQTVLSLQNSFSSEP 58

RESULT 9
KTHY_XANAC
ID KTHY_XANAC STANDARD; PRT; 227 AA.
AC Q8PF67;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thymidylate kinase (EC 2.7.4.9) (dTMP Kinase).
GN Name=tmk; OrderedLocusNames=XAC4014;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Phosphorylation of dTMP to form dTDP in both de novo and
CC salvage pathways of dTTP synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE012049; AAM38850.1;
CC HSSP; O05891; 1N5L.
CC HMAP; MF_00165; 1.
CC InterPro; IPR000062; Thymidylate kin.
CC Pfam; PF02223; Thymidylate kin; 1.
CC TIGRFAMS; TIGR00041; dTMP Kinase; 1.
CC PROSITE; PS01331; THYMIDYLATE KINASE; FALSE NEG.
KW ATP-binding; Complete proteome; Kinase; Nucleotide biosynthesis;
KW Transferase.
FT NP_BIND 16 23 ATP (Potential).
SQ SEQUENCE 227 AA; 24030 MW; 7C78722F1BCFA211 CRC64;

Query Match 7.2%; Score 169.5; DB 1; Length 227;
Best Local Similarity 27.4%; Pred. No. 0.00023;
Matches 58; Conservative 37; Mismatches 68; Indels 49; Gaps 8;

QY 254 VVAIEGLDAGTKTQVTSQVADSLKA-----VLKSPSPCIGQW-----RKLFDD 297
Db 11 LIAIEGIDGAGKTTLARLAATLDAAGARYVLSKEPTN--GPMGTQLRQSAATGRLSAE 68

QY 298 EPTIIRAFYSLGNYIVASEIAKESAKSPVIVDRYWHSTATYATVATVSVGLQHLRPAHH 357
Db 69 EASLLRDRHEHVDTLIAPALAGDI---VILDRYFSPWVAYQ-----GAGLGP---- 114

QY 358 PVYQWPEDL-----KPDLIILLTVSPERLQRLQGRMGKTRBEEAELEANSVFQKV 410
Db 115 -----LDELLERNAFAFRPDVILLLLDLPPTGLARIRARGDAPNHEETQDNL-----ERC 164

QY 411 EMSYQRMENPCGHVVDASPSREKVLQTVLSLI 442
Db 165 RTIFAALPLPGKHVDASADSDSVLRQAHAI 196

RESULT 10
KTHY_XANCP
ID KTHY_XANCP STANDARD; PRT; 227 AA.
AC Q8P3Y6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thymidylate kinase (EC 2.7.4.9) (dTMP Kinase).
GN Name=tmk; OrderedLocusNames=XCC3931;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).

```

```

CC -!- FUNCTION: Phosphorylation of dTMP to form dTDP in both de novo and
CC salvage pathways of dTTP synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE012513; AAM43153.1; -.
DR HSSP; P37345; STMP.
DR HAMAP; MF 00165; -.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; 1.
DR TIGRFAMs; TIGR00041; DTMP_Kinase; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; FALSE NEG.
KW ATP-binding; Complete proteome; Kinase; Nucleotide biosynthesis;
FT NP BIND 16 23 ATP (Potential).
SQ SEQUENCE 227 AA; 24000 MW; BDB0D3C9E66753C CRC64;

Query Match 7.0%; Score 166.5; DB 1; Length 227;
Best Local Similarity 26.8%; Pred. No. 0.00038;
Matches 57; Conservative 35; Mismatches 70; Indels 51; Gaps 8;

QY 254 VVAIEGLDATGKTTVTQSVADSL-----KAVLLKSPSCIGQW-----RKTFDD 297
DB 11 LIAIEGIDGAGKTTIARSALTLQAGARVVLVSKEPTN--GPWGTQLRQSAATGRLSAQD 68

QY 298 EPTTIIRAFYSLGNVIVASEITAKSAPSVIVDRYHSTATYATATEVSGGLQHLPPAHH 357
DB 69 EVDLLDRRHVEALIPALARGEI---VILDRYFPMWAY---QGAAGL----- 113

QY 358 PVYQWPEDDL-----KPDILLITVSPERLQGRGMEKTRAEAELEANSVPFQK 409
DB 114 -----PDALLAANDFAFRPDILLILLDLPPTGLARIGARDAPNHFTQNL-----ER 163

QY 410 VEMSYQRMENPGCHVDASPSREKVLQTVLSLI 442
DB 164 CRAIFPALQLPGKHVIDASADSVLRQAHAVV 196

RESULT 11
KTHY_XILFPU
ID KTHY_ARCFU STANDARD; PRT; 196 AA.
AC O30175;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE Probable thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN Name=tkk; OrderedLocustNames=AF0061;
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9889475; DOI=10.1038/37052;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,

```

```

RA Woese C.R., Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RA reducing archaeon Archaeoglobus fulgidus.";
RA Nature 390:364-370(1997).
CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001102; AAB91163.1; -.
DR PIR; B69257; E69257.
DR HSSP; P37345; STMP.
DR TIGR; AF0061; -.
DR HAMAP; MF 00165; -.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; 1.
DR TIGRFAMs; TIGR00041; DTMP_Kinase; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; 1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide biosynthesis;
FT NP BIND 7 14 ATP (Potential).
SQ SEQUENCE 196 AA; 22349 MW; F9CF20D740A5A367 CRC64;

Query Match 6.6%; Score 155; DB 1; Length 196;
Best Local Similarity 28.6%; Pred. No. 0.0022;
Matches 61; Conservative 37; Mismatches 75; Indels 40; Gaps 10;

QY 254 VVAIEGLDATGKTTVTQSVADSL-----KAVLLKSP-PSGIGQWRKIFDDEPTTIIRAFY 307
DB 2 LIAVEGIDGAGKTTIAAYIAELLKEGKYKVKVLPKPGDSKFG--KKIKSSEERLSPEBEL 59

QY 308 SLGNVIVASEI-AXES-----AKSPVIVDRYHSTATYATATEVSGGLQHLPPAHHPVY 360
DB 60 EL--FLKDRIDARENILPALQSGYAVVMDRYEYNATYQSGARGIDARL-----IR 108

QY 361 QWPEDDL-KPDILLITVSPERLQGRGMEKTRAEAELEANSVPFQKVMYSQRMEN 419
DB 109 EMNEKIAPKPDILTLLDVEPIALERVKRKGLSPFEKLD-----YLRKVRKCFLENAD 162

QY 420 PGCHVDASPSRE-----KVLQTVLSLIONS 445
DB 163 ETTVVVDASKPLEEVKKEVRKVIESFLNLKKN 195

RESULT 12
KTHY_XILFPA
ID KTHY_XILFPA STANDARD; PRT; 208 AA.
AC Q9PFS7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN Name=tkk; OrderedLocustNames=XF0580;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
RA Alvares R.H., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto C., Docena C., El-Dorri H.,

```



```

RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A., Ferro J.A.,
RA Praga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA da Souza A.P., Terenzi M.P., Truffi D., Tsai S.M., Tsuchako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Phosphorylation of dTMP to form dTDP in both de novo and
CC salvage pathways of dTTP synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003904; AAF83390.1; ALT_INIT.
DR HAMAP; MF_00165; -.
DR InterPro; IPR000062; Thymidylate kin.
DR Pfam; PF02223; Thymidylate kin; 1.
DR TIGRFAMs; TIGR00041; dTMP kinase; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; FALSE NEG.
KW ATP-binding; Complete proteome; Kinase; Nucleotide biosynthesis;
FT NP_BIND 7 14 ATP (Potential).
SQ SEQUENCE 208 AA; 22436 MW; 5563ADA4F9D541AE CRC64;
Query Match 6.5%; Score 154.5; DB 1; Length 208;
Best Local Similarity 27.7%; Pred. No. 0.0026;
Matches 59; Conservative 32; Mismatches 71; Indels 51; Gaps 9;
Qy 254 VVAIEGLDAGTKTTVTSQVADSLKAVLL-----KSPPCICGQWRKIF-----DD 297
Db 2 LVAIEGIDGAGKTTLARSALKRGVGLTWSKEPTN--GPWGMLLRQSAATGRFSPEE 59
Qy 298 EPTIIRAFYSLGNYIVASEIAKESAKSPVIVRYWHSTATYATATVSGGLQHLPPAHH 357
Db 60 EVDVLLDRDRQHVEDLIVPMIGRAV---VILDRYFPPSWAY-----QGAAGL----- 104
Qy 358 EPTIIRAFYSLGNYIVASEIAKESAKSPVIVRYWHSTATYATATVSGGLQHLPPAHH 357
Db 105 -----FVDALLEANAFAPRDPVLLLDVPPVIGLQRIWERG--STPNHFETTEN-----LSR 154
Qy 410 VEMSQRMPNPGCHVVDASPSREKVLQTVLSLI 442
Db 155 CRDIFLALELPSKRVIDATANAETVLSAALV 187
RESULT 13
ID_KTHY_XYLFIT STANDARD; PRT; 208 AA.
AC Q87B89;
DT 10-OCT-2003 (Rel. 42, Created)

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN Name:tmk; OrderedLocusNames=PD1569;
OS Xylella fastidiosa (strain Temecual / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorriy H., Tsai S.M.,
RA Carver H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuranae E.E.,
RA Marinho C.L., Gaglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Branco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.F., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- FUNCTION: Phosphorylation of dTMP to form dTDP in both de novo and
CC salvage pathways of dTTP synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE012559; AAO29411.1; -.
DR HSP; P37345; STMP.
DR HAMAP; MF_00165; -.
DR InterPro; IPR000062; Thymidylate kin.
DR Pfam; PF02223; Thymidylate kin; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; FALSE NEG.
KW ATP-binding; Complete proteome; Kinase; Nucleotide biosynthesis;
FT NP_BIND 7 14 ATP (Potential).
SQ SEQUENCE 208 AA; 22455 MW; 254E0C9268E33633 CRC64;
Query Match 6.3%; Score 148.5; DB 1; Length 208;
Best Local Similarity 27.2%; Pred. No. 0.0072;
Matches 58; Conservative 31; Mismatches 73; Indels 51; Gaps 9;
Qy 254 VVAIEGLDAGTKTTVTSQVADSLKAVLL-----KSPPCICGQWRKIF-----DD 297
Db 2 LVAIEGIDGAGKTTLARSALKRGVGLTWSKEPTN--GPWGMLLRQSAATGRFSPEE 59
Qy 298 EPTIIRAFYSLGNYIVASEIAKESAKSPVIVRYWHSTATYATATVSGGLQHLPPAHH 357
Db 60 EVDVLLDRDRQHVEDLIVPMIGRAV---VILDRYFPPSWAY-----QGAAGL----- 104
Qy 358 EPTIIRAFYSLGNYIVASEIAKESAKSPVIVRYWHSTATYATATVSGGLQHLPPAHH 357
Db 105 -----FVDALLEANAFAPRDPVLLLDVPPVIGLQRIWERG--STPNHFETTEN-----LSR 154
Qy 410 VEMSQRMPNPGCHVVDASPSREKVLQTVLSLI 442
Db 155 CRDIFLALELPSKRVIDATANAETVLSAALV 187

```

RESULT 14

Q6L016 PRELIMINARY; PRT; 202 AA.
 AC Q6L016;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative thymidylate kinase (EC 2.7.4.9).
 GN OrderedLocNames=PROL101;
 OS Schizosaccharomyces pombe.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Picrophilaceae; Picrophilus.
 OX NCBI_TaxID=92076;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 9790 / ATCC 700027;
 RX PubMed=15184674;
 RA Fuetterer O., Angelov A., Liesegang H., Gottschalk G., Schleper C.,
 RA Schepers B., Dock C., Antranikian G., Liebl W.;
 RT "Genome sequence of Picrophilus torridus and its implications for life
 around pH 0.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096 (2004).
 DR EMBL; AE017261; AA43686.1; -;
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0004798; F:thymidylate kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000062; Thymidylate_kin.
 DR Pfam; PF02223; Thymidylate_kin; 1.
 DR TIGRFAMs; TIGR00041; dTMP_kinase; 1.
 DR PROSITE; PS01331; THYMIDYLATE_KINASE; 1.
 KW Complete proteome; Kinase; Transferase.
 SQ SEQUENCE 202 AA; 23548 MW; ECD5A5FC43B67EB36 CRC64;

Query Match 6.1%; Score 145; DB 2; Length 202;
 Best Local Similarity 30.0%; Pred. No. 0.012;
 Matches 45; Conservative 22; Mismatches 71; Indels 12; Gaps 4;

QY 255 VAIEGLDATGKTTVTSQVADSLKAVLKSPSS--CIGQWRKIFDEPTIIRAFVSLGNY 312
 DB 9 IALGIDGSGKTLUAGDIASVTGYLTREPTDFCYDIADYDESSIIINFFLTIDRY 68
 QY 313 IVASEIAKESKSPVIVDRVWHSTATYATATEVSGGLQHLPPAHPVYQWPE-----LLK 368
 DB 69 MQHKEI-KNHLINGVSDRVVFSIAVQ-----GSGMEKRFKNWDETISWMLDVSRIIM 122
 QY 369 PDLILLTVSPERLQRLQGRGMKTRBA 398
 DB 123 PDLIIYLKIDPGIALKRLNLRKNEKNTDA 152

RESULT 15

KTHY_SCHPO STANDARD; PRT; 210 AA.
 AC P36590; 074528;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
 GN Name=tmpl; Synonyms=tmpl; ORFNames=SPCC70.07c;
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93003330; PubMed=13271149;
 RX Abaigar L.T., Yen Y.I., Jong A.Y.;
 RA "Functional and structural conservation of Schizosaccharomyces pombe
 dTMP kinase gene.";
 RT dTMP kinase gene.";
 RL Biochim. Biophys. Acta 1132:222-224 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne R., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson K.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Weijens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880 (2002).
 CC -!- FUNCTION: Catalyzes the conversion of dTMP to dTDP.
 CC -!- CATALYTIC ACTIVITY: ATP + dTMP = ADP + dTDP.
 CC -!- PATHWAY: Biosynthesis of dTMP from dTDP.
 CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X65868; CAA46698.1; -;
 DR EMBL; AL023794; CAA19357.1; -;
 DR PIR; S28955; S28955.
 DR PIR; T41553; T41553.
 DR HSP; P00572; ITMK.
 DR GeneDB Spombe; SPCC70.07c; -;
 DR InterPro; IPR000062; Thymidylate_kin.
 DR Pfam; PF02223; Thymidylate_kin; 1.
 DR TIGRFAMs; TIGR00041; dTMP_kinase; 1.
 DR PROSITE; PS01331; THYMIDYLATE_KINASE; 1.
 KW ATP-binding; Kinase; Nucleotide biosynthesis; Transferase.
 FT NP_BIND 14 21 ATP (Potential).
 FT CONFLICT 33 39 SQHEKAE -> LNMKRLK (in Ref. 1).
 FT CONFLICT 59 59 K -> T (in Ref. 1).
 FT CONFLICT 80 93 TIQIVYEQINKGV -> PSIIYRANQORCN (in Ref. 1).
 FT CONFLICT 125 125 P -> T (in Ref. 1).
 FT CONFLICT 164 164 F -> L (in Ref. 1).
 FT CONFLICT 186 186 S -> YA (in Ref. 1).
 FT CONFLICT 191 191 H -> D (in Ref. 1).
 SQ SEQUENCE 210 AA; 24249 MW; 4266144AEDAB68C0 CRC64;

Query Match 6.1%; Score 144.5; DB 1; Length 210;
 Best Local Similarity 26.1%; Pred. No. 0.014;
 Matches 58; Conservative 39; Mismatches 76; Indels 49; Gaps 11;
 QY 248 QKGRFQVVAIEGLDATGKTTVTSQVADSL-----KAVLKSPSSCIGQWRKIFD----- 296
 DB 3 KQNGRLIVIEGLDRSGKSTQCQLLVKLLSQHEKAELEFPDRTTAIGKIDYLVKESV 62
 QY 297 ---DEPTIIRAFVSLGNYIVASEIAKESKSPVIVDRVWHSTATYATATEVSGGLQHL 352

Db 63 QNDQ---VHLLFSANRWETIQYIEQINKGVTCILDY-----AFSGIAFSAKGLD--- 113
Qy 353 PPAHHEVYQW---PEDLLKPDILILLTVSPEERLQORGMKTRTEAELEANSVFRQ 408
Db 114 -----WEMCKSPDRGLPRPDLVIFLNVDP--RIAATRGQYGEERYEXIEMQ-----E 158
Qy 409 KVMSYORME----NPGCHVV--DASPSREKVLQTVLSLIQ 444
Db 159 KVLKNFQRLQKPEFREGLEFITLDASSSLEDVHSQIVDLVSN 200

Search completed: October 22, 2004, 23:28:54
Job time : 163 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 23:32:53 ; Search time 104 Seconds
(without alignments)
1397.766 Million cell updates/sec

Title: US-10-681-223-2
Perfect score: 2362
Sequence: 1 MAFARLLRGPLSGPLGRR.....SRKVLQTVLSLQNSFSEP 449

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pap:*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pap:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pap:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2362	100.0	449	14	US-10-277-032-2
2	2362	100.0	449	15	US-10-681-223-2
3	795.5	33.7	508	14	US-10-277-032-4
4	795.5	33.7	508	15	US-10-681-223-4
5	155	6.6	193	14	US-10-369-493-911
6	140	5.9	246	15	US-10-282-122A-67367
7	131.5	5.6	205	15	US-10-282-122A-53917
8	128.5	5.4	199	14	US-10-369-493-11256
9	128.5	5.4	253	16	US-10-437-963-196428
10	127.5	5.4	205	14	US-10-369-493-21577
11	127	5.4	259	14	US-10-369-493-22800
12	126.5	5.4	263	15	US-10-425-114-64336
13	126	5.3	210	9	US-09-815-242-11027

11	126	5.3	210	15	US-10-282-122A-58176	Sequence 58176, A
15	125.5	5.3	205	14	US-10-369-493-1320	Sequence 1320, Ap
16	125.5	5.3	205	14	US-10-369-493-20385	Sequence 20385, A
17	125	5.3	212	15	US-10-282-122A-61591	Sequence 61591, A
18	124.5	5.3	212	15	US-10-282-122A-77370	Sequence 77370, A
19	124.5	5.2	1664	16	US-10-437-963-107768	Sequence 107768, A
20	121	5.1	980	14	US-10-156-761-11338	Sequence 11338, A
21	120	5.1	206	15	US-10-282-122A-50167	Sequence 50167, A
22	120	5.1	212	14	US-10-369-493-23045	Sequence 23045, A
23	118.5	5.0	188	14	US-10-369-493-21453	Sequence 21453, A
24	118.5	5.0	216	14	US-10-369-493-22188	Sequence 22188, A
25	118.5	5.0	216	16	US-10-754-929-3	Sequence 3, Appli
26	118.5	5.0	1357	16	US-10-437-963-164009	Sequence 164009, A
27	118	5.0	193	14	US-10-369-493-18097	Sequence 18097, A
28	117.5	5.0	195	14	US-10-369-493-70	Sequence 70, Appl
29	117.5	5.0	213	9	US-09-815-242-10129	Sequence 10129, A
30	117.5	5.0	213	14	US-10-369-493-797	Sequence 797, App
31	117.5	5.0	213	15	US-10-282-122A-56520	Sequence 56520, A
32	117.5	5.0	213	16	US-10-754-929-2	Sequence 2, Appli
33	117.5	5.0	1750	16	US-10-437-963-194022	Sequence 194022, A
34	117	5.0	211	14	US-10-369-493-18331	Sequence 18331, A
35	116.5	4.9	213	15	US-10-282-122A-59458	Sequence 59458, A
36	116.5	4.9	218	14	US-10-369-493-5538	Sequence 5538, Ap
37	116	4.9	196	14	US-10-369-493-8559	Sequence 8559, Ap
38	115	4.9	206	15	US-10-282-122A-47929	Sequence 47929, A
39	114	4.8	1532	16	US-10-437-963-112468	Sequence 112468, A
40	113.5	4.8	210	14	US-10-369-493-286	Sequence 286, App
41	113	4.8	348	14	US-10-369-493-4190	Sequence 4190, Ap
42	112.5	4.8	201	14	US-10-369-493-4374	Sequence 4374, Ap
43	112.5	4.8	206	15	US-10-282-122A-49192	Sequence 49192, A
44	112.5	4.8	213	15	US-10-282-122A-55481	Sequence 55481, A
45	112.5	4.8	1753	16	US-10-437-963-107684	Sequence 107684, A

ALIGNMENTS

RESULT 1

```

US-10-277-032-2
; Sequence 2, Application US/10277032
; Publication No. US20030087294A1
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305 DIV
; CURRENT APPLICATION NUMBER: US/10/277,032
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 449
; TYPE: PRT
; ORGANISM: HomoSapien
US-10-277-032-2

```

Query Match	100.0%	Score	2362	DB	14	Length	449
Best Local Similarity	100.0%	Pred. No.	4.4e+201				
Matches	449	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MAFARLLRGPLSGPLGRRGVGCAGAMAPPCRFVLELPDCTLAHFALGADADADADPDP	60				
Db	1	MAFARLLRGPLSGPLGRRGVGCAGAMAPPCRFVLELPDCTLAHFALGADADADADPDP	60				
Qy	61	RLAALIGPERSYSLCVPTPDAGCGARVRAARLHQLRRGPFQRCOLLRLCYCP	120				
Db	61	RLAALIGPERSYSLCVPTPDAGCGARVRAARLHQLRRGPFQRCOLLRLCYCP	120				
Qy	121	GGGAGGAQGGFLLRDPLDPTDFTQALLLGGACQAPRHILGFEADPRGQLWRLEWQ	180				
Db	121	GGGAGGAQGGFLLRDPLDPTDFTQALLLGGACQAPRHILGFEADPRGQLWRLEWQ	180				

Db	121	GGQAGGACQGGFLRLDPLDDPTDTRQALLLELIGACQAPRPHLGEFADPRGQLWQRLMEVQ	180
Qy	181	DGRRLQVGCAQVWPVPPEPLHPVVPDLPSSVFFDREAAARAVLBECTSFIEARAVLDLV	240
Db	181	DGRRLQVGCAQVWPVPPEPLHPVVPDLPSSVFFDREAAARAVLBECTSFIEARAVLDLV	240
Qy	241	DQCPKQIQKGFQVVAIEGLDATCKTIVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPT	300
Db	241	DQCPKQIQKGFQVVAIEGLDATCKTIVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPT	300
Qy	301	IIRAFYSLGNVIYASBIAXESAKSPVIVDRYWHSTATYALATEVSGGLQHLPPAHFPVY	360
Db	301	IIRAFYSLGNVIYASBIAXESAKSPVIVDRYWHSTATYALATEVSGGLQHLPPAHFPVY	360
Qy	361	QWPEDLKPOLILILLITVSPERIORLQGRGMEKTRERAELEANSVFQKVMYSORMENP	420
Db	361	QWPEDLKPDUIILLITVSPERIORLQGRGMEKTRERAELEANSVFQKVMYSORMENP	420
Qy	421	GCHVVDAASPSEKVLQTVLSLIONSFSPEP	449
Db	421	GCHVVDAASPSEKVLQTVLSLIONSFSPEP	449

```

RESULT 2
US-10-681-223-2
; Sequence 2, Application US/10681223
; Publication No. US20040081999A1
; GENERAL INFORMATION:
; APPLICANT: Midg-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0601305 DIV.-II
; CURRENT APPLICATION NUMBER: US/10/681,223
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 10/277,032
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-681-223-2

```

Query Match	100.0%	Score 2362;	DB 15;	Length 449;
Best Local Similarity	100.0%;	Pred. No. 4.4e-201;		
Matches 449;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAFARLRGLPSGLGRGV	CAGAMAPCCRFVLELP	PDCTLAHFALCAGADGADADPD 60
DB	1	MAFARLRGLPSGLGRGV	CAGAMAPCCRFVLELP	PDCTLAHFALCAGADGADADPD 60
QY	61	RLAALLGPPERSYSLCV	PVTPDAGCGAVRAARL	HQRLHQLRRGPFQRCOLLILCYCP 120
DB	61	RLAALLGPPERSYSLCV	PVTPDAGCGAVRAARL	HQRLHQLRRGPFQRCOLLILCYCP 120
QY	121	GQAGGAQGFLLRPDL	DDPTRQALLLGLCACQAP	RPHLGEFADPRGQLWQRLNEVQ 180
DB	121	GQAGGAQGFLLRPDL	DDPTRQALLLGLCACQAP	RPHLGEFADPRGQLWQRLNEVQ 180
QY	181	DGRRILQVCAQVPVPE	PLHPVDPDLSSVFPD	REAAARVILECTSFIFEARAVLDLV 240
DB	181	DGRRILQVCAQVPVPE	PLHPVDPDLSSVFPD	REAAARVILECTSFIFEARAVLDLV 240
QY	241	DCQCPKQIQKGKQFWA	IIIGLDAATGKTTVTSQ	VADSLKAVLAKSPSCICQWRKIFDDEPT 300
DB	241	DCQCPKQIQKGKQFWA	IIIGLDAATGKTTVTSQ	VADSLKAVLAKSPSCICQWRKIFDDEPT 300
QY	301	IIRRAFYSIGNVINVA	SEAKSAPFVIDRYYW	HSATYATIAETVSGGLQHLPPAHFPVY 360

Db	301	IIIRAFYSLGNIVYVASEIAKESAKSPVIDRVYWHSTATVATIEVSGGLQLPAPHVY	360
Qy	361	QWPEDDLKPDLLILLTVSPBEERLORLQGRGMKTRFEAELEANSVFRQKVEVSYQRMENP	420
Db	361	QWPEDDLKPDLLILLTVSPBEERLORLQGRGMKTRFEAELEANSVFRQKVEVSYQRMENP	420
Qy	421	GCHVVDASPSREKVLQTVLSLIQNSFSEP	449
Db	421	GCHVVDASPSREKVLQTVLSLIQNSFSEP	449

```

RESULT 3
US-10-277-032-4
; Sequence 4, Application US/10277032
; Publication No. US2003087294A1
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305 DIV
; CURRENT APPLICATION NUMBER: US/10/277,032
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 508
; TYPE: PRT
; ORGANISM: HomoSapien
US-10-277-032-4

```

Query Match	33.7%;	Score	795.5;	DB	14;	Length	508;
Best Local Similarity	48.1%;	Pred. No.	9.8e-62;				
Matches	177;	Conservative	0;	Mismatches	4;	Indels	187;
Gaps							2;
QY	225	ECTSFIPEARAVLDLVDQCPKQIQKGFQVAIEGLD	ATGKTTVTQSVADSLKAVLLKSP	284			
Db	1	ECTSFIPEARAVLDLVDQCPKQIQKGFQVAIEGLD	ATGKTTVTQSVADSLKAVLLKSP	60			
QY	285	-----	-----	284			
Db	61	CTSFPEARAVLDLVDQCPKQKGFQVAIEGLD	ATGKTTTQQCTSFPEARAVLDLVDQC	120			
QY	285	-----	-----	303			
Db	121	PKVEQKGFQVAIEGLD	ATGKTTLTQHFKSLSLSSYRHPSCIGQWRKIFDDEPTIIR	180			
QY	304	RAFYSLGNYIVASIIAKESAKS	-----	325			
Db	191	RAFYSLGNYIVASIIAKESAKSPVIVDRYHSTAT	YPCIKENYVASEIAKESPVIVDRYM	240			
QY	326	-----	-----	337			
Db	241	HSTATYPPCIKPVBEDLIMNLLSFEFFILWANYLVASEIAKES	TNFPVIVDRYHSTAT	300			
QY	338	TYATATEVSGGLQHLPPAHHPVYQWPEDLIKPDILLITVS	PEERLQRLQGRGMKEKTREE	397			
Db	301	TYATATEVSGGLQHLPPAHHPVYQWPEDLIKPDILLITVS	PEERLQRLQGRGMKEKTREE	360			
QY	398	AELEANSV	405				
Db	361	AEAIATEV	368				

RESULT 4
US-10-681-223-4
; Sequence 4, Application US/10681223
; Publication No. US20040081999A1
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

```

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305 DIV-II
; CURRENT APPLICATION NUMBER: US/10/681,223
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 10/277,032
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-681-223-4

Query Match 33.7%; Score 795.5; DB 15; Length 508;
Best Local Similarity 48.1%; Pred.No. 9.8e-62;
Matches 177; Conservative 0; Mismatches 4; Indels 187; Gaps 2;

QY 225 ECTSFIPEARAVLDLVDQCPKQIQKGKFOVAIEGLDATGKTTVTSQSVADSLKAVLLKSP 284
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 ECTSFIPEARAVLDLVDQCPKQIQKGKFOVAIEGLDATGKTTVTSQSVADSLKAVLLKSP 60

QY 285 ----- 284
Db 61 CTSFIPEARAVLDLVDQCPKQKGKFOVAIEGLDATGKTTTQCTSFIPPEARAVLDLVDQC 120

QY 285 -----PSCIGQWRKIFDDEPTIIR 303
Db 121 PREVQKGFQVIAIEGLDATGKTTLTQHFKSLRSLSSYSRHPSCIGQWRKIFDDEPTIIR 180

QY 304 RAFYSLGNIVASEIAKESAKS----- 325
Db 181 RAFYSLGNIVASEIAKESAKSPVIVDRVWHSTATPYCIKFNVASEIAKESPVIVDRYW 240

QY 326 -----PVIIVRYWHSTA 337
Db 241 HSTATYPPCIKPVEEDLLMMNLLSPEEPFILWANVIVASEIAKESTNFPVIVRYWHSTA 300

QY 338 TYAIAIEVSGGLQHLPPAHHPVYQWPEDLLKPDLLLLLTVSPEERLQRLQGRGMEKTREE 397
Db 301 TYAIAIEVSGGLQHLPPAHHPVYQWPEDLLKPDLLLLLTVSPEERLQRLQGRGMEKTREE 360

QY 398 AELEANSV 405
Db 361 ABAIAIEV 368

RESULT 5
US-10-369-493-911
; Sequence 911, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 911
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Archaeoglobus fulgidus
US-10-369-493-911

```

Best Local Similarity 27.6%; Pred. No. 0.00066;
Matches 59; Conservative 32; Mismatches 89; Indels 34; Gaps 10;

QY 250 GKFOVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPTIIRAFYS----- 308
Db 2 IIAIEGIDAGKNTIVSAIKERFDADVIGFP-----RYEQSIHAKLAQALYGSMDL 54
QY 309 -----LCNYIVASEIAKESAKSPVIVDRYWHSTATYATVSG 347
Db 55 TDSAYAMATLFDALDRYDAKAVLGRYVGTSKV-----VLLDRYVASNAAYSAAR----- 102
QY 348 GLOHLPAPHPVYQWPEL-----LKPDILLLTVSPEERLQRLQ-----GRQWEK 393
Db 103 -----TRNDAMVQWQVELEFEELGLPVPDIHLLHTSPELAAQRAORREATDASRKLD 156
QY 394 TREAELEANSVERQKVMESYQRMENPGCHV--VDASPS--REKVLQTV 438
Db 157 YEEDAGLOERT--FOAYESLAQKQWMSPLLVHPDESDPTVTQRIQAL 203

RESULT 8
US-10-369-493-11256
; Sequence 11256, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11256
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Methanosarcina mazei
US-10-369-493-11256

Query Match 5.4%; Score 128.5; DB 14; Length 199;
Best Local Similarity 26.0%; Pred. No. 0.0052;
Matches 58; Conservative 39; Mismatches 71; Indels 55; Gaps 12;

QY 249 KGKFOVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPT-----I 301
Db 1 RKG--LITLEGIDGSGKSTVAEK-----LQKNPE--IKAFKPVFTRETRGTLTGDA 48
QY 302 IRRAPYSLGN-----YIVASEIAKESAK-----SPVIVDRYWHSTATYATVSG 347
Db 49 VEKAIQSDTDQFAELFLTADHAELAKLIKPALENGKIVISDRSYAQ----- 101
QY 348 GLOHLPAPHPVYQWPEL-----LKPDILLLTVSPEERLQRLQGRQWEKTRAEALEA 402
Db 102 GITLKTRENLPL--EWVKDLHRSWTIVPDLTFLDIRPISIERCCKRQSKFKLE--- 157
QY 403 NSVFRQKVMESYQRM--ENPGCHV--DASPSREKVLQTVLSLI 442
Db 158 ---FLOGVRAIFLKLAAADDPFVVIDASRSPFXIEKEVVKKI 197

RESULT 9
US-10-437-963-196428
; Sequence 196428, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

Best Local Similarity 27.6%; Pred. No. 0.00066;
Matches 59; Conservative 32; Mismatches 89; Indels 34; Gaps 10;

QY 250 GKFOVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPTIIRAFYS----- 308
Db 2 IIAIEGIDAGKNTIVSAIKERFDADVIGFP-----RYEQSIHAKLAQALYGSMDL 54
QY 309 -----LCNYIVASEIAKESAKSPVIVDRYWHSTATYATVSG 347
Db 55 TDSAYAMATLFDALDRYDAKAVLGRYVGTSKV-----VLLDRYVASNAAYSAAR----- 102
QY 348 GLOHLPAPHPVYQWPEL-----LKPDILLLTVSPEERLQRLQ-----GRQWEK 393
Db 103 -----TRNDAMVQWQVELEFEELGLPVPDIHLLHTSPELAAQRAORREATDASRKLD 156
QY 394 TREAELEANSVERQKVMESYQRMENPGCHV--VDASPS--REKVLQTV 438
Db 157 YEEDAGLOERT--FOAYESLAQKQWMSPLLVHPDESDPTVTQRIQAL 203

RESULT 7
US-10-282-122A-53917
; Sequence 53917, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53917
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53917

Query Match 5.6%; Score 131.5; DB 15; Length 205;
Best Local Similarity 22.4%; Pred. No. 0.0029;
Matches 51; Conservative 39; Mismatches 69; Indels 69; Gaps 9;


```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64336
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3632-055-C3_FLI.pep
US-10-425-114-64336

Query Match      5.4%; Score 126.5; DB 15; Length 263;
Best Local Similarity 25.5%; Pred. No. 0.012;
Matches 60; Conservative 34; Mismatches 70; Indels 71; Gaps 13;

QY 250 GKQVVAIEGLDGTGTTTQSVADSIKAVLLKSPSPCIGQWKIFDDPTIIRAFYSL 309
Db 59 GRGALVVLGELDRSGTKSQCARLLSLFKGYNRAE-----GWR--FPDRAT-----SV 104

QY 310 GNYIVASEIAKES-----AKSPFVIVDRYWHSTATYATA 342
Db 105 GQ-WISSYLANDSQLDRTTHILFSANRWKRALMESKLLSGTTLIVDRYSYS---GVA 159

QY 343 TEVSGGLQHLPPAHHPVYQW---PED-LLKPDILLITVSPERLQRLQGRGMKTEKEEA 398
Db 160 FSAKAGLD-----IEWCKAPENGLIADIVILDVQPEKAAER-GYGGGER-YEKI 208

QY 399 ELEANSYFQKVMYSQRMENPGCHVYDA-----SPSREKVLQTVLSLQNSFSEP 449
Db 209 E-----FQKVAHVHSLRSTWVKVVDGFLPMETVEEKLRLDLATSCIQCQGNP 257

RESULT 13
US-09-815-242-11027
; Sequence 11027, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11027
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11027

Query Match      5.3%; Score 126; DB 9; Length 210;
Best Local Similarity 26.1%; Pred. No. 0.0094;
Matches 57; Conservative 32; Mismatches 91; Indels 38; Gaps 10;

QY 249 KGKQVVAIEGLDGTGTTTQSVADSL-----KAVLLKSPSPC-----IGQWRKIFDD 297
Db 2 KGKF--IVIEGLEGAGKSAHQSVVRVLHELGIQDVVFTREPQGTPLAEKRLHLIKHETE 59

QY 298 EPTIIRAFYSLGNYIVASEIAKESAKSP-----VIVDRYWHSTATYATAIEVSGGLQH 351
Db 60 EPVTDKAEILLML--YAARIQLVENVIKPALMOCKQWVVGDRHDMSSQAY-----QGGGRQ 111

QY 352 LPPAHHPVYQWPEDLI---KPDILLITVSPERLQRLQGRGMKTEKEEAELANSVFRQ 408
Db 112 LDP--HFMLTLKHTVLGNFEPDLTIYLDIDPSVGLARARGELDRIEQMDLD----FFH 165

QY 409 KVMSYQRM--ENPGCHVVDASPSREKVLQTVLSLQIN 444
Db 166 RTRARYLELVKONPKAVINAEQSIELVQADIESAVKN 203

RESULT 14
US-10-282-122A-58176
; Sequence 58176, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58176
```

```
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-10-282-122A-58176

Query Match      5.3%; Score 126; DB 15; Length 210;
Best Local Similarity 26.1%; Pred. No. 0.0094;
Matches 57; Conservative 32; Mismatches 91; Indels 38; Gaps 10;

QY 249 KGKQVVAIEGLDATGKTTVTQSVADSL-----KAVLLKSPSC-----IGQWRKIFDD 297
Db 2 KGKF--IVIEGLEGAGKSSAHQSVRVVHELGIQDVVFTREPGGTPLAEKLRHLIKHETE 59
QY 298 EPTIIRAFYSLGNIYIVASEIAKESAKSP-----VIVDRYWHSTATYAIATEVSGGLQH 351
Db 60 EPTVDKAEALLML--YAARIQLVENVIKPALMQGKWVVGDRHDMSSQAY-----QGGRQ 111
QY 352 LPFAHPVYQWPEDDL---KPDLLILLTVSPERLQRLQGRGMEKTEEAELANSVFRQ 408
Db 112 LDP--HPMLTKETVLGNFEFDITVLDIDPSVGLARARGELDRIEQMDLD----FFH 165
QY 409 KVMSYQRM--ENPGCHVDVADSPREKVLQTVLSLIQN 444
Db 166 RTRARYLELVKDNPKAVVNAEQSIELVQADIESAVKN 203

RESULT 15
US-10-369-493-1320
; Sequence 1320, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkie, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1320
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-1320

Query Match      5.3%; Score 125.5; DB 14; Length 205;
Best Local Similarity 23.3%; Pred. No. 0.01;
Matches 53; Conservative 44; Mismatches 71; Indels 59; Gaps 10;

QY 249 KGKQVVAIEGLDATGKTTVTQSVAD-----SLKAVLLKSP-PSICIGQW-RKIFDDEPTI 301
Db 2 RGYF--IVLEGIDGSGKTTQAKLIAEFWDKGYEVLTKTEPTDSELGKLIRRIIEESVI 59
QY 302 I-BRAFYSLGNIVYASEIAKESAK-----SPVIVDRYWHSTATYAIATEVSGGLQH 351
Db 60 DGSKISYEAEALLFAADRAEHVKIILPALSEGKVVICDRFYSSL----- 105
QY 352 LPFAHPVYQWPEDDL-----LKPDLILLTVSPERLQRLQGRGMEKTEEAEL 399
Db 106 -----AYQWARGLDLNLQVNSEFAPRPDLAILDLPVKESLRRIKRG-----TLTE 153
QY 400 LEANSVFRQVMSYQRMEN--PCCHVVDASPSREKVLQTVLSLIQN 444
Db 154 FDKIVELQKVRHNYLKAEMFPEMRIVNALSSIEDIHSDIVALVKH 200
```

Search completed: October 22, 2004, 23:40:34
Job time : 107 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 23:05:48 ; Search time 29 Seconds
(without alignments)
1026.786 Million cell updates/sec

Title: US-10-681-223-2

Perfect score: 2362

Sequence: 1 MAFARLLRGPLSGPLGRR.....SREKVIQTVLSLQNSFSEP 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2362	100.0	449	4	US-09-984-880-2
2	2362	100.0	449	4	US-10-277-032-2
3	795.5	33.7	508	4	US-09-984-880-4
4	795.5	33.7	508	4	US-10-277-032-4
5	129.5	5.5	936	4	US-09-252-991A-28002
6	118.5	5.0	216	4	US-09-632-553-3
7	118	5.0	691	4	US-09-252-991A-25690
8	117.5	5.0	213	4	US-09-632-553-2
9	117.5	5.0	219	4	US-09-489-039A-7540
10	109	4.6	955	4	US-09-252-991A-24254
11	108.5	4.6	635	4	US-09-252-991A-25642
12	107	4.5	212	3	US-09-259-109-2
13	107	4.5	212	4	US-09-583-110-5141
14	107	4.5	2294	4	US-09-252-991A-17231
15	106	4.5	1427	4	US-09-252-991A-20577
16	105.5	4.5	204	3	US-09-134-001C-3311
17	105.5	4.5	388	4	US-09-252-991A-21572
18	105.5	4.5	564	4	US-09-252-991A-25356
19	105	4.4	444	4	US-09-252-991A-27505
20	104.5	4.4	1665	4	US-09-858-664A-2
21	104.5	4.4	1665	4	US-10-274-978-2
22	104	4.4	494	4	US-09-252-991A-23320
23	104	4.4	1067	4	US-09-252-991A-30526
24	104	4.4	1544	3	US-09-413-814-46
25	103.5	4.4	582	4	US-09-252-991A-25366
26	103.5	4.4	751	4	US-10-020-079-8
27	103.5	4.4	764	4	US-10-020-079-6

28 103.5 4.4 854 4 US-10-020-079-4 Sequence 4, Appli
29 103.5 4.4 870 4 US-10-020-079-2 Sequence 2, Appli
30 101.5 4.3 414 4 US-09-252-991A-30034 Sequence 30034, A
31 101.5 4.3 1343 4 US-09-171-991-2 Sequence 2, Appli
32 101 4.3 337 4 US-09-252-991A-31851 Sequence 31851, A
33 100.5 4.3 305 4 US-09-252-991A-20357 Sequence 20357, A
34 100.5 4.3 422 4 US-09-252-991A-30625 Sequence 30625, A
35 100 4.2 280 4 US-09-252-991A-32145 Sequence 32145, A
36 100 4.2 605 4 US-09-252-991A-25512 Sequence 25512, A
37 99.5 4.2 227 4 US-09-543-681A-7769 Sequence 7769, Ap
38 99.5 4.2 435 4 US-09-252-991A-24702 Sequence 24702, A
39 99.5 4.2 507 4 US-09-252-991A-17308 Sequence 17308, A
40 99.5 4.2 657 4 US-09-252-991A-28001 Sequence 28001, A
41 99.5 4.2 776 4 US-10-020-079-24 Sequence 24, Appl
42 99.5 4.2 789 4 US-10-020-079-22 Sequence 22, Appl
43 99.5 4.2 889 4 US-10-020-079-20 Sequence 20, Appl
44 99.5 4.2 895 4 US-10-020-079-18 Sequence 18, Appl
45 99.5 4.2 2285 4 US-09-252-991A-17790 Sequence 17790, A

ALIGNMENTS

RESULT 1

US-09-984-880-2

; Sequence 2, Application US/09984880

; Patent No. 6489153

; GENERAL INFORMATION:

; APPLICANT: Ming-Hui WEI

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: C1001305

; CURRENT APPLICATION NUMBER: US/09/984,880

; CURRENT FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 449

; TYPE: PRT

; ORGANISM: HomoSapien

; US-09-984-880-2

Query Match 100.0%; Score 2362; DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 2.2e-232;

Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAFARLLRGPLSGPLGRRGVCAGAMAPCRFVLELPDCTLAHFAFGADAPGDADAPDP	60
Db	1	MAFARLLRGPLSGPLGRRGVCAGAMAPCRFVLELPDCTLAHFAFGADAPGDADAPDP	60
Qy	61	RLAALLGPPERSYSLCVPTPDAGCGARVRAARLHQRLLHQLRRGPFQRCQLRLCYCP	120
Db	61	RLAALLGPPERSYSLCVPTPDAGCGARVRAARLHQRLLHQLRRGPFQRCQLRLCYCP	120
Qy	121	GGAGGAQQGFLRDPLDDPTDTRQALLELIGACQEAAPRPHLGFEDAPRQQLWQRLWEVQ	180
Db	121	GGAGGAQQGFLRDPLDDPTDTRQALLELIGACQEAAPRPHLGFEDAPRQQLWQRLWEVQ	180
Qy	181	DGRRLQVGCQAVVPPEPLHPVVPDLPGSVVPPDRFAARVLEECTSFTEPEARAVLDLV	240
Db	181	DGRRLQVGCQAVVPPEPLHPVVPDLPGSVVPPDRFAARVLEECTSFTEPEARAVLDLV	240
Qy	241	DQCPKIQKGFQVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPT	300
Db	241	DQCPKIQKGFQVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPT	300
Qy	301	IIRRAFYSIGNYIVASEIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHFVY	360
Db	301	IIRRAFYSIGNYIVASEIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHFVY	360
Qy	361	QWPEDLKPDLLILLTVSPEERLORLOGRMKTRTEALEANSVROKVMESYQRMENP	420
Db	361	QWPEDLKPDLLILLTVSPEERLORLOGRMKTRTEALEANSVROKVMESYQRMENP	420

Db 361 QWPEDLKPDLLILLTVSPEERLQRLQGRMGKTRAEAELEANSVFRQKVMYSQRMENP 420

QY 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

Db 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

RESULT 2

US-10-277-032-2

; Sequence 2, Application US/10277032

; Patent No. 6664087

; GENERAL INFORMATION:

; APPLICANT: Ming-Hui WEI

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001305 DIV

; CURRENT APPLICATION NUMBER: US/10/277,032

; CURRENT FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/984,880

; PRIOR FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 449

; TYPE: PRT

; ORGANISM: HomoSapien

US-10-277-032-2

Query Match 100.0%; Score 2362; DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 2.2e-232; Indels 0; Gaps 0;

Matches 449; Conservative 0; Mismatches 0;

QY 1 MAFARRLLRGPLSGPLGRRGVCAGAMAPCRFVLELPDCTLAHFALGADAPGDADPDP 60

Db 1 MAFARRLLRGPLSGPLGRRGVCAGAMAPCRFVLELPDCTLAHFALGADAPGDADPDP 60

QY 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHORLLHQIRGPGRCOLLRLCYCP 120

Db 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHORLLHQIRGPGRCOLLRLCYCP 120

QY 121 GQAGGAGQGLLRLDPTDPTQALLELGCACAPRPHLGEFEADPRGQLWQRLWEVQ 180

Db 121 GQAGGAGQGLLRLDPTDPTQALLELGCACAPRPHLGEFEADPRGQLWQRLWEVQ 180

QY 181 DGRRLQVCAQVVPPEPLHPVDPDLPSSVFPDREARAVLEECTSFIFEARAVLDLV 240

Db 181 DGRRLQVCAQVVPPEPLHPVDPDLPSSVFPDREARAVLEECTSFIFEARAVLDLV 240

QY 241 DQCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVILKSPSCIGQWRKIFDDEPT 300

Db 241 DQCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVILKSPSCIGQWRKIFDDEPT 300

QY 301 IIRAFYSLGNIVVASEIAKESAKSPVIVDRYWHSTATYATVATEVSGGLQHLPPAHHPVY 360

Db 301 IIRAFYSLGNIVVASEIAKESAKSPVIVDRYWHSTATYATVATEVSGGLQHLPPAHHPVY 360

QY 361 QWPEDLKPDLLILLTVSPEERLQRLQGRMGKTRAEAELEANSVFRQKVMYSQRMENP 420

Db 361 QWPEDLKPDLLILLTVSPEERLQRLQGRMGKTRAEAELEANSVFRQKVMYSQRMENP 420

QY 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

Db 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

RESULT 3

US-09-984-880-4

; Sequence 4, Application US/09984880

; Patent No. 6489153

; GENERAL INFORMATION:

; APPLICANT: Ming-Hui WEI

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001305

; CURRENT APPLICATION NUMBER: US/09/984,880

; CURRENT FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 508

; TYPE: PRT

; ORGANISM: HomoSapien

US-09-984-880-4

Query Match 33.7%; Score 795.5; DB 4; Length 508;

Best Local Similarity 48.1%; Pred. No. 2.7e-72;

Matches 177; Conservative 0; Mismatches 4; Indels 187; Gaps 2;

QY 225 ECTSFIFEARAVLDLVQCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVILKSP 284

Db 1 ECTSFIFEARAVLDLVQCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVILKSP 60

QY 285 ----- 284

Db 61 CTSFIFEARAVLDLVQCPKQIQKGFQVVAIEGLDATGKTTTQCTSFIFEARAVLDLVQDC 120

QY 285 ----- PSCIGQWRKIFDDEPTIIR 303

Db 121 PKEVQKGFQVVAIEGLDATGKTTLTQHFKSLSLSSYSRHPSHCIGQWRKIFDDEPTIIR 180

QY 304 RAFYSLGNIVVASEIAKESAKS----- 325

Db 181 RAFYSLGNIVVASEIAKESAKSPVIVDRYWHSTATYPCIKFENYVASEIAKESPVIVDRYW 240

QY 326 ----- PVIIVDRYWHSTA 337

Db 241 HSTATYPCIKPVEEDLLMMNLLSFEPPFIWANYLVASEIAKESKSTNFPVIVDRYWHSTA 300

QY 338 TYATATEVSGGLQHLPPAHHPVYQWPEDLKPDLLILLTVSPEERLQRLQGRMGKTRTEE 397

Db 301 TYATATEVSGGLQHLPPAHHPVYQWPEDLKPDLLILLTVSPEERLQRLQGRMGKTRTEE 360

QY 398 AELEANSV 405

Db 361 AEATATEV 368

RESULT 4

US-10-277-032-4

; Sequence 4, Application US/10277032

; Patent No. 6664087

; GENERAL INFORMATION:

; APPLICANT: Ming-Hui WEI

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001305 DIV

; CURRENT APPLICATION NUMBER: US/10/277,032

; CURRENT FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/984,880

; PRIOR FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 508

; TYPE: PRT

; ORGANISM: HomoSapien

US-10-277-032-4

Query Match 33.7%; Score 795.5; DB 4; Length 508;

Best Local Similarity 48.1%; Pred. No. 2.7e-72;

Matches 177; Conservative 0; Mismatches 4; Indels 187; Gaps 2;

QY 225 ECTSFIFEARAVLDLVQCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVILKSP 284

Db 1 ECTSFIPPEARAVLDVDCPKQIQKGFQVVAIEGLDATGKTTVTSVADSLKAVLLKSP 60
QY 285 ----- 284
Db 61 CTSFIPPEARAVLDVDCPKQKGFQVVAIEGLDATGKTTTQCTSFIPPEARAVLDVDC 120
QY 285 ----- PSCIGQWRKIFDDEPTIR 303
Db 121 PKEVQKGFQVIAIEGLDATGKTTTQHFHKSLSRLSSYRHPSCIGQWRKIFDDEPTIR 180
QY 304 RAYFSLGNYIVASEIAKESAKS ----- 325
Db 181 RAYFSLGNYIVASEIAKESAKSPVIVDRYWHSTATYPCIKENYVASEIAKESPVIVDRY 240
QY 326 ----- PVIVDRYWHSTA 337
Db 241 HSTATYPCIKPVEEDLLMMNLSFEPFFILWANYLVASEIAKESNTFPVIVDRYWHSTA 300
QY 338 TYAIAIEVSGLOHLPAPHHVYQWPEDLKPDLLILLITVSPERLORLOGRGMEKTREE 397
Db 301 TYAIAIEVSGLOHLPAPHHVYQWPEDLKPDLLILLITVSPERLORLOGRGMEKTREE 360
QY 398 AELEANSV 405
Db 361 AEAIATEV 368
RESULT 5
US-09-252-991A-28002
; Sequence 28002, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28002
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28002
Query Match 5.8%; Score 129.5; DB 4; Length 956;
Best Local Similarity 22.4%; Pred. No. 0.00076;
Matches 124; Conservative 55; Mismatches 182; Indels 193; Gaps 29;
QY 24 AGAMAPPCCRF-----VLELPDCTLAHFALGADA-----PGDADAP----- 58
Db 250 AGAVPRLRAADPAALGDLRLQDPDPALLHGQLAALAPGKRPDTQCGEPDLDVPLRRHP 309
QY 59 ---DPLA---ALLGPP-----ERSYSLCVPTPDAGCGARVRAARLHORLLHQLRRGP 106
Db 310 RRDAPGLARRPPAQCGPPAGGDSRRRRVAVHAPRPQPPAALAGVRLRRLLHRRRAVD 369
QY 107 FORCOLRLCYCGGQAG---GAQQGFLRLDPLDPPDFTQALL-----ELLGACQEA 156
Db 370 PER--LRQQLLPFGFTHRLHAGRGPLRFD--RPTVRQPAAGHAHIRGEHLHACGDS 425
QY 157 PRHLGEFEADPRGQLWQRLWEVQDGRRLVGCAQVV-----PVPEPPLHPV 203
Db 426 RRARRA---TDFPGLARRFGAA--GKRRAGRGILSLHRLPLGYSGGFFRRPQEP 480
QY 204 VPDLP-----SVVFPDREA-----ARA 221
Db 481 MNDADSLQYQVRQQAIRSLFEIVGSSSEGTIVVDREARIVWINKRYAQRFGGLADPAQA 540

QY 222 VLEECTSFIPPEA-----RAVIDLVDCPKQIQKGFQVVAIEGLDATGKT--TVT 269
Db 541 IGRPCSEVIFGSLRELVGSGKPIILLDMDSA-----KDFLVVMRLFIHDDAGEVIGAIG 595
QY 270 QSVADSLKAVLLKSP--PSCIGQWRKIFDDEPTI--IRRAFYSYLSGNYIVAS-----EIAKES 322
Db 596 FALLDEBSL---SPLVTRYLSMRELASTRSOLQARQAKYSFHSFVGTSFASLEVKRRA 652
QY 323 AKSPVIVDRYWHSTATYAIATEVSGGLOHL-----PPAHPVYQ-----WPEDLLKP 369
Db 653 RRGAA-----SEAPVLLLGTTGCKELLAHAIHSASPAHAKAFVSVNAAAIPESILEA 705
QY 370 DLILLITVSP-----BERLQ-----LQGRGMEKTREAELE 401
Db 706 E---LFGTAPGFTGAERKRPGLQVQAQGTFLDDEIGDMPALQOK--LLRVLQKEBE 761
QY 402 A---NSVFRQKQVEM 412
Db 762 AVGSNEVIRSDIRL 775
RESULT 6
US-09-632-553-3
; Sequence 3, Application US/09632553
; Patent No. 6689595
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Crystallization and Structure Determination of
; TITLE OF INVENTION: Staphylococcus Aureus Thymidylate Kinase
; FILE REFERENCE: 6245.NCP
; CURRENT APPLICATION NUMBER: US/09/632,553
; CURRENT FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/147,117
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-632-553-3
Query Match 5.0%; Score 118.5; DB 4; Length 216;
Best Local Similarity 25.0%; Pred. No. 0.00097;
Matches 54; Conservative 37; Mismatches 80; Indels 45; Gaps 11;
QY 250 GKQVVAIEGLDATGKTTVTSVADSLK--AVLLKSP-----SCIGQWRKIFDDEPTI 301
Db 3 GRGKLILIEGLDRTGKTTQCNILYKKLPNCKLLKPERSTRIGGLINEY--LTDSDFOL 60
QY 302 IIRA---FYSLGNYIVASEIAKESAKSP-VIVDRYWHSTATYAIATEVSGGLOHLPAAH 357
Db 61 SDQAIHLLFSANRWEIVDKIKOLLEGKNIVMDRYVYGVAYSAAKGTNG----- 110
QY 358 PVYQW---PE-DLLKPDLLILLITVSPERLORLOGRGMEKTREAELEANSVPROKEMS 413
Db 111 MLDLWCLQPDVGLLKPDLTLFLSTQVDNNAEKSGFGDERYE-----TVKFOEVKQT 163
QY 414 YQ-----RMENPGCHVVDASPSREKVLQTVLSLI 442
Db 164 FMKLLDKERKGDSEITIVDVT---NGKIQEVEALI 196
RESULT 7
US-09-252-991A-25690
; Sequence 25690, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

```

US-09-632-553-2
Query Match      5.0%; Score 117.5; DB 4; Length 213;
Best Local Similarity 23.6%; Pred. No. 0.0012;
Matches 52; Conservative 42; Mismatches 85; Indels 41; Gaps 10;

QY      255 VAIEGLDAGTKTTVQSVADSLK-----AVLLKSP-----SCIGQWRK 293
DB      6 IVIEGLEGAGKTTAENVVVTLEQLGIRDMVFTPEPGTQLAEKLSIWLDIKSVGD--E 63

QY      294 IFDEPTTIRRAFYSLGNYIVASSETAKSASP-VIVDRYWHSTATYAIAIEVSGLOHL 352
DB      64 VITDKAEVL--MFYAARVQLVETVTKPALANGTWVIGDRHDLSTQAY-----QGGGRGI 115

QY      353 PPAHPHPVYQWPDLL--KPDILLLTVSPEERLQGRG-MEKTREEAELEANSVPRQ 408
DB      116 D--QHMIALTRDAVLGDFRPDUTLYDVTPEVGLKRAARGELDRIEQESDFFNRTAR 173

QY      409 KVMESYQRMENPGCHVVDVADSPSRKVLQTVLSLQNSFSF 448
DB      174 YLELAAQ--DKSIHTIDATQPLEAVMDAIRTTVTHWVKE 210

RESULT 9
US-09-489-039A-7540
; Sequence 7540, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7540
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7540

Query Match      5.0%; Score 117.5; DB 4; Length 219;
Best Local Similarity 22.1%; Pred. No. 0.0013;
Matches 52; Conservative 41; Mismatches 85; Indels 57; Gaps 8;

QY      248 QKGFQVVAIEGDATGKTTVQSVADSLK-----VLLKGPSPSCI----- 288
DB      5 KKMRSNVIVIEGLEGAGKTTARQLVVTETQSAGIHDVFTREPLGTILAEKLSIWLDIQ 64

QY      289 GOWRKIFDDEPTTIRRAFYSLGNYIVASSETAKSASP-VIVDRYWHSTATY----- 339
DB      65 STGDVINDKAEVL--MFYAARVQLVETVTKPALARGQVWVIGDRHDLSTQAYQGGGRGID 122

QY      340 -----AIAIEVSGLOHLPPAHPHPVYQWPDLLKPDILLLTVSPEERLQGRG-MEK 393
DB      123 RTMIATLRDAVLGD-----FRPNLTLYLDVTPEVGLQRAARGELDR 164

QY      394 TREAELEANSVPRQKVMESYQRMENPGCHVVDVADSPSRKVLQTVLSLQNSFSF 448
DB      165 IEQESMNFNRTARVYLELA--AADPSRTVDATQPLDAVARDIRATIAQWMAE 216

RESULT 10
US-09-252-991A-24254
; Sequence 24254, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

```

QY 340 -----ATATSEVSGLOHLPAAHPVYQWPEDLIKPOLILLITVSPERLQRCGRG-MEX 393
 Db 123 RTMIATLRDAVLGD-----PPNLTLYLDVTPVGLQRAFAREGELDR 164
 QY 394 TREAEALEANSVFROKVEYSQYRMENPGCHVDPASPSREKVLQTVLSLQNSFSE 448
 Db 165 IEQSMNFFNRTARYILELA--AADPSITVDATQPLDAVARDIRATIAQWMAE 216

RESULT 10
 US-09-252-991A-24254
 ; Sequence 24254, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136


```

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24254
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24254

Query Match
Best Local Similarity 28.1%; Pred. No. 0.094;
Matches 74; Conservative 14; Mismatches 83; Indels 92; Gaps 15;

QY 4 ARLLRG-----PLSGPLG-----REGVCA-----GAMAPCRFVLELPDCT 41
DB 706 ARRAVEGAGDPPSPVSRPLLAGPTIDRGLAAGRPAGDSVQGAAGGTGGLASDPGAT 765
QY 42 LAHFALGAD---APGDADAPDP-----RLAALLG-----PPERSY 73
DB 766 DHGAGLGCGEPGAPGPA-APLPRQRLRMVAGVAGNLSHGKRTALPRPGTQPEHRAI 824
QY 74 SLICVPVTPDAGCA-----RVRAARLHQRL---HQLRRGPQRCQLLRLCYCPGQAG 125
DB 825 PRHAP-GPLGAGAGRRPPRAQPVRLHLRLGGGHCRRQ-----PGQPG 869
QY 126 GAAQGFLLRDPLDDPTQRLLELLGACQFAPRPHLGEFADPRGQLWORLDMEVQDGRLL 185
DB 870 HGRFP-----AQPRPQ-----SPERPRPRPGEL---PRPTVQVLGQAQRPV 912
QY 186 QVGCAGQWVPVPEPLHPVVPDLP 208
DB 913 RQDRSASLPERQLRP-DGJLP 934

RESULT 11
US-09-252-991A-25642
; Sequence 25642, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25642
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25642

Query Match
Best Local Similarity 21.6%; Pred. No. 0.053;
Matches 98; Conservative 49; Mismatches 170; Indels 137; Gaps 20;

QY 6 RLLRGPLSGPLGRGVCAGAMAPCRFVLELPDCTLAHFALGADAGDADDPRLAAL 65
DB 280 RLAAHHHPGTADRHMTCKNATPE-----LLEL-----SPRLDADPGVRLALI 325
QY 66 ----LGPERSYSLCVPTPDAGCARVRAARLHQLRLHQLRRGPQRCQLRLCYCPG 121
DB 326 ELADLELPEALPLLIARLGRDPPGVGEARLLE-----AWEEDAVVDALCAA-- 374
QY 122 GQAGGAQGGFLRLDPLDDPTQRLLELLGACQF-APR---PHLGEFADPRGQLWRL 176

```

```

DB 375 -----LADPV--PAVADAAQSLGELKEPAAGRLLPLWGHADAFVRASVLRAL 421
QY 177 WEVQDGRRLQVCAQWVPVPEPLHPVVPDLPSSVVFFDREARAVLEECTSFIPEARAV 236
DB 422 REL-----RLESAA-----PALAALGDPQAA-----VRREAVAV 451
QY 237 LDLDVQCPKQIQKGFQVVAIEGLDATGKTTVTQSVADSLKAVLLKSPESCI---GOWR 292
DB 452 LGWLRHQFALAEAR---LASADVDFEVRRATGALGSLREATVL--PALCAALADAQWQ 506
QY 293 KIFDDEPTIIRAFYSLGNYIVASEIAKESAPV---IVDRYHSTATYATATVSGGL 349
DB 507 -----VREEAATLG-----KLGREEAGEPLKALADDYQVRLR---AARALGEL 549
QY 350 QHLPPAHHPVYQWPEDLKPDILLITVSPERLORLOGRMEKTEBAALEANSVFROK 409
DB 550 RHRP-----AREALEALLGHPIGNLRKEAALALGELADPA 584
QY 410 VEMSYORMENPGCHVVVDASPSREKVLQTVLSLIQ 443
DB 585 SAQALRVABG-----DGDPEVRKAVRIALAQLR 612

RESULT 12
US-09-259-109-2
; Sequence 2, Application US/09259109
; Patent No. 6270762
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K. R.
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Biswas, Sanjoy
; APPLICANT: Chaliker, Allison F.
; APPLICANT: Ingraham, Karen A.
; APPLICANT: Traini, Christopher M.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: tdk
; FILE REFERENCE: GM10201
; CURRENT APPLICATION NUMBER: US/09/259,109
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-259-109-2

Query Match
Best Local Similarity 4.5%; Score 107; DB 3; Length 212;
Matches 52; Conservative 34; Mismatches 90; Indels 36; Gaps 8;

QY 254 VVAIEGLDATGKTTVTQSVADSLK---AVLLKSPSCIGQWRKIFD-----DEP 299
DB 6 LVSLGEGAGKTSVLEALLPILKEKGEVLTTRPQGVVLGKIREVILDPSTQMDAK 65
QY 300 TIIRAFYSLGNYIVASEIAKESAPVIVDRYHSTATY-----AIATEVSGGLQHLPP 354
DB 66 TELLYIASRRQHLVEKVLFALEAGKLVIMDRFIDSVAYQGGFGRGLDTEADWLN--- 121
QY 355 AHPVVQWPEDLKPDILLITVSPERLORLOGRMEKTEBAALEANSV-FRQKVEMS 413
DB 122 -----QFATDGLKPDLTLYFDIEVEGLARI---HANSREVNRLDLEGLDLHKVRQ 172
QY 414 Y-QRMENPGCHVV--DASPSREKVLQTVLSLI 442
DB 173 YLSLLDKEGNRIVKIDASLPLEQVETTRAVL 204

RESULT 13
US-09-583-110-5141
; Sequence 5141, Application US/09583110
; Patent No. 6699703

```

```

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5141
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5141

Query Match          4.5%; Score 107; DB 4; Length 212;
Best Local Similarity 24.3%; Pred. No. 0.014;
Matches 52; Conservative 34; Mismatches 90; Indels 36; Gaps 8;

QY 254 VVAIEGLDAGTKTTVTSVADSLK---AVLLKSPSCIGQWRKIFD-----DEP 299
DB 6 LVSLGEGGAGKTSVLEALLPILEKCVETLTTEPGGVLLGKIREVILDPHTQMDAK 65
QY 300 TIIRAFYISGNVIVASEIAKESAKSPVIVDRYVWHSTATY-----AIATVSGGLQHLPP 354
DB 66 TELLVYIARQHLVEKVLPALEAGKLVIMDRFIDSSVAYQGFGRGLDIEADWLN---121
QY 355 AHPVNVQWPEDLKPLDILLTTSPERLQIQRGMEKTRERAELEANSV-FRQVEMS 413
DB 122 -----QFATDGKPLDTLYFDIEVEEGLARI---AANSREVNRLDLEGLDLHKVROG 172
QY 414 Y-ORMENPGCHVY--DASPEREKVLOTVLSLI 442
DB 173 YLSLKEGNNRIVKIDASLEPLEQVVTXAVL 204

RESULT 14
US-09-252-991A-17231
; Sequence 17231, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17231
; LENGTH: 2294
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17231

Query Match          4.5%; Score 107; DB 4; Length 2294;
Best Local Similarity 27.7%; Pred. No. 0.6;
Matches 65; Conservative 16; Mismatches 90; Indels 64; Gaps 11;

QY 5 RRLIRGLSGPLIGRR-----GVCAGAMAPPCRFVLELPDCTLAHFALGADAPGADA 57
DB 1477 RRLHQGLDQRLGTQPAARPLAGTAGAARARRVAR-----TTA--AVGAPGPGRG 1529
QY 58 PDRFLAALLGPPERSVSLCVPTFTDAGC--GARVRAARLHORLLHQRGPFQRCOLLRL 115

```

```

DB 1530 PAPPVAAATAGGA-----AGLRPQAAEAAGRRARLPHQRRREP-----1557
QY 116 LCYCPGQAGG--AQOGFLLRDLDPDTRQALLELLGACQEAP--RPHLGEFEADPRGQ 171
DB 1568 -----RRAAGVRAQBRQAHRP-----AQRLHLPLRLSRGVPHRQPGQGRHHR--RGT 1613
QY 172 LWORLWEVDGRRLQVGCACQVVPVPEPFLHPVVDPLPSSVVFVDRBAARAVLBEC 226
DB 1614 VGARARPERCRRRQPGPRA-----SPVLPGLPAAVGRPARPYRSADHQC 1660

RESULT 15
US-09-252-991A-20577
; Sequence 20577, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20577
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20577

Query Match          4.5%; Score 106; DB 4; Length 1427;
Best Local Similarity 22.7%; Pred. No. 0.36;
Matches 113; Conservative 46; Mismatches 189; Indels 150; Gaps 24;

QY 4 ARRLRGLSGPLIGRRGVCAGAMAPPCRFVLELPDCTLAHFALG-----ADAPGADA 57
DB 773 ARRAVREP-----GR-----PADHFDHRAAVDLRRAAAGAGURDEHLYADRPFGPDR 821
QY 58 PD-----PRLAALIGP--PERSYSLCVPTFTDAGC-----GARVRAARLHORLL 99
DB 822 PDQOARHPHGRVQRQTAPRAPPRGDPARRADPPAAGADDHRGNGLRPGAAALRQR--879
QY 100 HOLRRGPFQRCOLLRL-----LCYCPGQAGGACQOGFLLRDLDD-----139
DB 880 -RRRQPLRPGRGDRDLRAGRHLHPVRAARLYPAGAQRGSRQ-----EPAQFATGRG 933
QY 140 -----PDTRQALLLEL-----LGACQEAP--RP-----HLGEFEADPR 169
DB 934 RSAGEQGMKRSYFNLSRLALAVGTCLACSVGPDYQRPQSPFPVASEHLGEFSGERR 993
QY 170 GQLWRLWEVDGRRLQVGCACQVVPVPEPFLHPVVDPLPSSVVFVDRBAAR--AVLEPCTS 228
DB 994 EAPW--WSFFEDD-----POLVRLVDQALARNHDIRARA 1025
QY 229 FIEEASAVLD--LVDQCPKQIQKGFQVVAIEGLDATKTTVTSQVADSLKAVLLKSPPS 286
DB 1026 NLESARALEFDDRWLDDQLPQVTSQAGYSRSIEQQLDYDGP--RRRLAESTRA-----G 1076
QY 287 CIGQWRKIFDEPTIIRRAFYSISGNIVASEIAKESAKSPVIVDRYVWHSTATVAIATEVS 346
DB 1077 FDAQWE--ID-----LFGRLGRLSDAALAEAAADADLRLVLSIAADTARAYFEIQ 1126
QY 347 GGLQHLPPAHPHYVQWPEDL-LKPDILLITTSPEERLQRLQGRGMEKTRERAELE--AN 403
DB 1127 GYQRLDVARAQVRSNRDTLELTRSLQLGSLGPEP-----VENAQANLLRSEAAIPPLAT 1182
QY 404 SVFRQKVMESYQRMENPG 421
DB 1183 ALSSARYRLDVLGRGAPG 1200

```

Search completed: October 22, 2004, 23:33:21
Job time : 32 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 20:24:38 ; Search time 116 Seconds
(without alignments)
1388.530 Million cell updates/sec

Title: US-10-681-223-2

Perfect score: 2362
Sequence: 1 MAFARLLRGLPGILGRR.....SREKVLQTVLSLQNSFSEP 449

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358720299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: _geneseqp1980s:*
- 2: _geneseqp1990s:*
- 3: _geneseqp2000s:*
- 4: _geneseqp2001s:*
- 5: _geneseqp2002s:*
- 6: _geneseqp2003as:*
- 7: _geneseqp2003bs:*
- 8: _geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2362	100.0	449	6	ABU08678 Novel hum
2	2362	100.0	449	7	ADCE3370 Human nov
3	2362	100.0	449	8	ADN49356 Human kin
4	2350	99.5	449	7	AEE38613 Human thy
5	2350	99.5	449	8	ADI40851 Human kin
6	2350	99.5	449	8	ADP84606 Human bre
7	2177.5	92.2	422	7	ADC39136 Novel hum
8	2177.5	92.2	422	8	ADH42559 Novel hum
9	1552	65.7	395	7	AEE38614 Mouse thy
10	1552	65.7	395	8	ADJ76250 Marker ge
11	1185	50.2	251	7	ADC39138 Novel hum
12	1185	50.2	251	8	ADH42565 Novel hum
13	1137	48.1	290	7	ADF76971 Novel hum
14	982.5	41.6	200	8	ADH42561 Novel hum
15	982.5	41.6	200	8	ADH42563 Novel hum
16	795.5	33.7	508	7	ADCE3372 Mouse thy
17	795.5	33.7	508	8	ADN49358 Mouse thy
18	677.5	28.7	431	5	ABB57297 Mouse isc
19	140	5.9	243	6	ABU39443 Protein e
20	137.5	5.8	283	3	AAG20877 Arabidops
21	132	5.6	208	8	ADN47368 Thermoco
22	131.5	5.6	205	6	ABU25993 Protein e
23	129.5	5.5	956	7	ABO79256 Pseudomon
24	128.5	5.4	224	3	AAG20878 Arabidops
25	128.5	5.4	224	3	AAG50881 Arabidops

26	128.5	5.4	263	3	AAG50880 Arabidops
27	127.5	5.4	205	4	AAB96152 Putative
28	126	5.3	210	2	AAY28787 E.coli th
29	126	5.3	210	4	AAY28787 Haemophil
30	126	5.3	210	4	AAY28787 Haemophil
31	126	5.3	210	7	ABU30252 Protein e
32	126	5.3	210	7	ADJ62328 H.influen
33	125	5.3	212	6	ABU33667 Protein e
34	124.5	5.3	212	6	ABU49446 Protein e
35	123.5	5.2	227	2	AAR07563 Polypepti
36	120	5.1	210	2	AAY28793 Schizosac
37	120	5.1	212	2	ABU22243 Protein e
38	119.5	5.1	204	2	AAY28785 Thymidyla
39	118.5	5.0	188	2	AAY28789 Thymidyla
40	118.5	5.0	216	4	AAY28796 Thymidyla
41	118.5	5.0	216	4	AAY28796 Thymidyla
42	118	5.0	210	7	ADJ62330 H.influen
43	118	5.0	681	7	ABO76944 Pseudomon
44	117.5	5.0	213	2	AAY28786 E.coli th
45	117.5	5.0	213	4	AAY28786 E.coli th

ALIGNMENTS

RESULT 1
ABU08678
ID ABU08678 standard; protein; 449 AA.
XX
AC ABU08678;
XX
DT 10-JUN-2003 (first entry)
XX
DE Novel human thymidylate kinase subfamily kinase.
XX
KW Human; thymidylate kinase subfamily; kinase; therapeutic agent;
KW immune response; kinase associated disorder; tissue typing;
KW Pharmacogenomic analysis; enzyme.
XX
OS Homo sapiens.
XX
FH Key
FT Modified-site 17..20 Location/Qualifiers
FT Modified-site /note="Amidation site"
FT Modified-site 21..26 /note="Myristoylation site"
FT Modified-site 121..126 /note="Myristoylation site"
FT Modified-site 181..184 /note="Amidation site"
FT Modified-site 259..266 /note="ATP/GTP-binding site motif A (P-loop)"
FT Modified-site 259..264 /note="Myristoylation site"
FT Modified-site 263..265 /note="Myristoylation site"
FT Modified-site /note="Protein kinase C phosphorylation site"
FT Modified-site 271..274 /note="Casein kinase II phosphorylation site"
FT Modified-site 275..277 /note="Protein kinase C phosphorylation site"
FT Modified-site 322..324 /note="Protein kinase C phosphorylation site"
FT Modified-site 378..381 /note="Casein kinase II phosphorylation site"
FT Modified-site 390..395 /note="Casein kinase II phosphorylation site"
FT Modified-site 394..397 /note="Myristoylation site"
FT Modified-site /note="Casein kinase II phosphorylation site"
FT Modified-site 407..414 /note="Tyrosine kinase phosphorylation site"
FT Modified-site 445..448 /note="Casein kinase II phosphorylation site"
XX

PN US6489153-B1.
XX 03-DEC-2002.
XX 31-OCT-2001; 2001US-00984880.
XX 31-OCT-2001; 2001US-00984880.
XX (PEKE) PE CORP NY.
XX Wei M, Ketchum KA, Beasley EM, Difrancesco V;
XX WPI; 2003-327246/31.
XX N-PSDB; ABX93989.
XX New human kinase polypeptides and nucleic acid molecules, useful in the
XX development of human therapeutic targets, or for diagnosing or treating a
XX disorder associated with the aberrant expression of the protein or
XX nucleic acid molecule.
XX Claim 1; Fig 2A-C; 49pp; English.
XX The invention describes a new isolated nucleic acid molecule encoding a
XX kinase of the thymidylate kinase subfamily. The human kinase polypeptides
XX and nucleic acid molecules are useful as models for the development of
XX human therapeutic targets, aid in the identification of therapeutic
XX proteins, or serve as targets for the development of human therapeutic
XX agents that modulate kinase activity in cells and tissues that express
XX the kinase. The kinase proteins are also useful in raising antibodies or
XX eliciting another immune response, as reagents in assays designed to
XX quantitatively determine levels of the protein in biological fluids, as
XX markers for tissues in which the corresponding protein is preferentially
XX expressed, or in screening a compound for the ability to stimulate or
XX inhibit interaction between the kinase protein and a molecule that
XX normally interacts with the kinase protein. The kinase proteins and
XX nucleic acid molecules can also be used in providing a target for
XX diagnosing a disease or a predisposition to a disease mediated by the
XX polypeptide or nucleic acid molecule, or for treating a disorder
XX associated with the absence of, inappropriate, or unwanted expression of
XX the protein or nucleic acid molecule. The antibodies are useful for
XX tissue typing, in isolating or detecting kinase proteins, or in
XX pharmacogenomic analysis. This is the amino acid sequence of the novel
XX human thymidylate kinase subfamily kinase
XX Sequence 449 AA;
Query Match 100.0%; Score 2362; DB 6; Length 449;
Best Local Similarity 100.0%; Pred. NO. 1.2e-230;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFARLLRGSLGRLGRGVCAGAMAPCRFVLELPDCTLAHFAIGADAPGDADPDP 60
DB 1 MAFARLLRGSLGRLGRGVCAGAMAPCRFVLELPDCTLAHFAIGADAPGDADPDP 60
QY 61 RLAAALGPPRSVSLCVPTVPDAGCGARVRAARLHORLLHOLRGPPQRCQLLRLLCYCP 120
DB 61 RLAAALGPPRSVSLCVPTVPDAGCGARVRAARLHORLLHOLRGPPQRCQLLRLLCYCP 120
QY 121 GGAGGAGQOQLLRDLPDPTQALLELIGACQAPRPHLGEFADPRGOLWRLWEVQ 180
DB 121 GGAGGAGQOQLLRDLPDPTQALLELIGACQAPRPHLGEFADPRGOLWRLWEVQ 180
QY 181 DGRRLQGVCAQVVPPEPLHPVVDLPSSVVFDPREAAARVLEECTSFPEARAVLDLV 240
DB 181 DGRRLQGVCAQVVPPEPLHPVVDLPSSVVFDPREAAARVLEECTSFPEARAVLDLV 240
QY 241 DQCFKQIQKGKQFQVVALEGLDATGKTTVTSQVADSLKXAVLLKSPSCIGQWRKIFDDEPT 300
DB 241 DQCFKQIQKGKQFQVVALEGLDATGKTTVTSQVADSLKXAVLLKSPSCIGQWRKIFDDEPT 300
QY 301 IIRAFYSLGNYIVASIIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHFPVY 360
DB 301 IIRAFYSLGNYIVASIIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHFPVY 360

QY 361 QWPEDLLKPDLLILLTVSPERLQRLQGRGMKTRREAELEANSVFRQKVMYSQRMENP 420
DB 361 QWPEDLLKPDLLILLTVSPERLQRLQGRGMKTRREAELEANSVFRQKVMYSQRMENP 420
QY 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
DB 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
RESULT 2
ADG63370
ID ADG63370 standard; protein; 449 AA.
XX AC ADC63370;
XX 19-DEC-2003 (first entry)
XX Human novel kinase protein.
XX Human; enzyme; kinase; chromosome 2.
XX Homo sapiens.
XX US2003087294-A1.
XX 08-MAY-2003.
XX 22-OCT-2002; 2002US-00277032.
XX 31-OCT-2001; 2001US-00984880.
XX (APPL-) APPLERA CORP.
XX Wei M, Ketchum KA, Beasley EM, Difrancesco V;
XX WPI; 2003-765435/72.
XX N-PSDB; ADC63369, ADC63371.
XX Novel isolated human kinase protein useful for drug screening assays, as
XX a target for diagnosing disease, pharmacogenomic analysis, and for
XX identifying compounds that modulate kinase activity.
XX Claim 1; SEQ ID NO 2; 48pp; English.
XX The invention relates to an isolated peptide consisting of or comprising
XX an amino acid sequence chosen from an amino acid sequence of human kinase
XX protein (ADC63370), an allelic variant of the kinase, orthologue or
XX fragment. Also included are an isolated antibody that selectively binds
XX to the kinase, an isolated nucleic acid molecule encoding the kinase,
XX proteins detailed above (appearing as the cDNA, ADC63369, and gene,
XX ADC63371), a gene chip comprising the nucleic acids, a transgenic non-
XX human animal comprising the nucleic acids; a nucleic acid vector
XX comprising the nucleic acids, a host cell containing the vector,
XX detecting the presence of the kinase in a sample, detecting the presence
XX of the nucleic acids in a sample, a pharmaceutical composition comprising
XX a kinase binding agent, treating a disease or condition mediated by a
XX human kinase protein and identifying a modulator of the expression of the
XX kinase involves. The kinase is useful for identifying a modulator, for
XX identifying an agent that binds to the kinase, for biological assays
XX related to kinases, drug screening assays, identifying compounds that
XX modulate kinase activity, to screen a compound for the ability to
XX stimulate or inhibit interaction between the kinase protein and a
XX molecule that normally interacts with kinase protein, for competitive
XX binding assays to discover compounds that interact with kinase,
XX pharmacogenomic analysis and treating a disorder characterised by an
XX absence of inappropriate or unwanted expression of the protein, useful as
XX target spot diagnosing active protein activity, disease or predisposition
XX to disease. The nucleic acid is useful for assessing expression in
XX disease states, diagnostic tools as an immunological marker for aberrant
XX protein for inhibiting protein function, tissue typing. The nucleic acids
XX are useful for facilitating protein trafficking or facilitate
XX manipulation of a protein for assay or production, probes, primers,

CC chemical intermediates, constructing recombinant vectors, expressing
CC antigenic portions of the proteins, for determining chromosomal positions
CC of the nucleic acid molecules, for making vectors containing the gene
CC regulatory regions, as hybridisation probes for determining the presence,
CC level, form and distribution of nucleic acid expression. The gene
CC encoding the kinase is located on human chromosome 2. The present
CC sequence represents the novel kinase.
XX
SQ Sequence 449 AA;
Query Match 100.0%; Score 2362; DB 7; Length 449;
Best Local Similarity 100.0%; Pred. No. 1.2e-230;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFARLLRGPLSGPLGRGVCAGAMAPPCFVLELPDCTLAHFALGADADGADADPP 60
Db 1 MAFARLLRGPLSGPLGRGVCAGAMAPPCFVLELPDCTLAHFALGADADGADADPP 60
QY 61 RLAAALGPPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRLCYCP 120
Db 61 RLAAALGPPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRLCYCP 120
QY 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
Db 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
QY 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
Db 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
QY 181 DGRRLQVCAQVVPPEPLHPVDPDLSSVVPDREAAVLEECTSFIPPEARAVLDIV 240
Db 181 DGRRLQVCAQVVPPEPLHPVDPDLSSVVPDREAAVLEECTSFIPPEARAVLDIV 240
QY 181 DGRRLQVCAQVVPPEPLHPVDPDLSSVVPDREAAVLEECTSFIPPEARAVLDIV 240
Db 181 DGRRLQVCAQVVPPEPLHPVDPDLSSVVPDREAAVLEECTSFIPPEARAVLDIV 240
QY 241 DQCPKQIQKGKQVVAIEGLDGTGKTVTQSVADSLKAVLLKSPSCICQWRKIFDEPT 300
Db 241 DQCPKQIQKGKQVVAIEGLDGTGKTVTQSVADSLKAVLLKSPSCICQWRKIFDEPT 300
QY 301 IIRRAFYSIGNIVASEIAKESAKSPVIVDRYVHSTATVAIATEVSGGLQHLPPAHHPYV 360
Db 301 IIRRAFYSIGNIVASEIAKESAKSPVIVDRYVHSTATVAIATEVSGGLQHLPPAHHPYV 360
QY 361 QWPELLKPDLLILLTVSPERLQRLQGRGMKTRTEAELEANSVFRQKVMYSQRMENP 420
Db 361 QWPELLKPDLLILLTVSPERLQRLQGRGMKTRTEAELEANSVFRQKVMYSQRMENP 420
QY 421 GCHVVDASPSREKVLTQVLSLQNSFSEP 449
Db 421 GCHVVDASPSREKVLTQVLSLQNSFSEP 449
RESULT 3
ADN49356
ID ADN49356 standard; protein; 449 AA.
XX AC ADN49356;
XX DT 15-JUL-2004 (first entry)
XX DE Human kinase protein.
XX KW Kinase; immune response; protein therapy; human; enzyme.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
FT Modified-site 17..20 /note= "Amidation site"
FT FT /note= "Amidation site"
FT Modified-site 21..26 /note= "N-myristoylation site"
FT FT /note= "N-myristoylation site"
FT Modified-site 121..126 /note= "N-myristoylation site"
FT FT /note= "N-myristoylation site"
FT Modified-site 181..184 /note= "Amidation site"
FT FT /note= "Amidation site"
FT Binding-site 259..266 /note= "ATP/GTP-binding site motif A"
FT FT /note= "ATP/GTP-binding site motif A"
FT Modified-site 259..264

FT Modified-site /note= "N-myristoylation site"
FT 263..285 /note= "protein kinase C phosphorylation site"
FT 271..274 /note= "Casein kinase II phosphorylation site"
FT 274..294 /note= "Helix 1"
FT 275..277 /note= "protein kinase C phosphorylation site"
FT 322..324 /note= "protein kinase C phosphorylation site"
FT 378..381 /note= "Casein kinase II phosphorylation site"
FT 390..395 /note= "N-myristoylation site"
FT 394..397 /note= "Casein kinase II phosphorylation site"
FT 407..414 /note= "Tyrosine kinase phosphorylation site"
FT 445..448 /note= "Casein kinase II phosphorylation site"
XX US2004081999-A1.
XX 29-APR-2004.
XX 09-OCT-2003; 2003US-00681223.
XX 31-OCT-2001; 2001US-00984880.
XX 22-OCT-2002; 2002US-00277032.
XX (APPL-) APPLERA CORP.
XX Wei M, Ketchum KA, Beasley EM, Di Francesco V;
XX WPI: 2004-347669/32.
XX N-PSDB; ADN49355, ADN49357.
XX New isolated human kinase proteins, useful for diagnosing or treating
XX disorders having an absence of, inappropriate, or unwanted expression of
XX the protein.
XX Claim 1; SEQ ID NO 2; 49pp; English.
XX The present invention provides human kinase polypeptides and the encoding
XX polynucleotides. The invention is useful in the development of human
XX therapeutic and diagnostic compositions, to identify therapeutic proteins
XX and serve as targets for the development of human therapeutic agents that
XX modulate kinase activity in cells and tissues, to elicit an immune
XX response and for diagnosing and treating disorders characterised by the
XX absence of inappropriate and unwanted expression of the protein. The
XX polypeptide is also useful in protein therapy. The present sequence is
XX the human kinase protein of the invention.
XX Sequence 449 AA;
Query Match 100.0%; Score 2362; DB 8; Length 449;
Best Local Similarity 100.0%; Pred. No. 1.2e-230;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFARLLRGPLSGPLGRGVCAGAMAPPCFVLELPDCTLAHFALGADADGADADPP 60
Db 1 MAFARLLRGPLSGPLGRGVCAGAMAPPCFVLELPDCTLAHFALGADADGADADPP 60
QY 61 RLAAALGPPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRLCYCP 120
Db 61 RLAAALGPPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRLCYCP 120
QY 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
Db 121 GGAGGAQGGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGQLWQRLWEYQ 180
QY 181 DGRRLQVCAQVVPPEPLHPVDPDLSSVVPDREAAVLEECTSFIPPEARAVLDIV 240

```

Db      181 DGRRLQVGCAQVVPPEPPLHPVDPDLPSSVVPDREAAARVLEECTSFIPARAVLDLV 240
QY      241 DQCPKQIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCICQWRKIFDEPT 300
Db      241 DQCPKQIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCICQWRKIFDEPT 300
QY      301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATVAIATEVSGGLQHLPPAHPVY 360
Db      301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATVAIATEVSGGLQHLPPAHPVY 360
QY      361 QWPEDLKPDLLILLTVSPEERLQRLQGRGMKTRREEALEANSVFRQKVMESYQRMENP 420
Db      361 QWPEDLKPDLLILLTVSPEERLQRLQGRGMKTRREEALEANSVFRQKVMESYQRMENP 420
QY      421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
Db      421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

RESULT 4
AAE38613
ID      AAE38613 standard; protein; 449 AA.
AC      AAE38613;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Human thymidylate kinase protein.
XX
KW      Human; thymidylate kinase; EC 2.7.4.9; ATP:dtMP phosphotransferase; TK;
KW      enzyme; cancer; gene therapy; chromosome 2p25.2.
XX
OS      Homo sapiens.
XX
PN      WO2003064642-A1.
XX
PD      07-AUG-2003.
XX
PF      27-JAN-2003; 2003WO-EP000783.
XX
PR      28-JAN-2002; 2002US-0351424P.
PR      25-NOV-2002; 2002US-0428711P.
XX
PA      (FARB ) BAYER AG.
XX
PI      Liou J;
XX
DR      WPI; 2003-646151/61.
DR      N-PSDB; AAD58542.
XX
PT      New polynucleotide and its encoded thymidylate kinase (TK), useful for
PT      identifying modulators of TK activity (e.g. agonists or antagonists) that
PT      provides therapeutic effects, and in gene therapy for treating cancers.
XX
PS      Claim 1; Fig 2; 11opp; English.
XX
CC      The invention relates to human thymidylate kinase (TK) polypeptides (EC
CC      2.7.4.9) and polynucleotides. TK is also referred to as ATP:dtMP
CC      phosphotransferase. TK sequences are useful for identifying test
CC      compounds, that may act as agonists or antagonists at the receptor site
CC      and which can be regulated to provide therapeutic effects. Vectors
CC      comprising TK polynucleotides are useful in the preparation of a
CC      medicament for modulating the activity of TK in a disease, particularly
CC      cancer. The invention is also useful in gene therapy. The present
CC      sequence is human thymidylate kinase protein. TK gene is located on
CC      chromosome 2p25.2
XX
SQ      Sequence 449 AA;

```

Query Match 99.5%; Score 2350; DB 7; Length 449;
 Best Local Similarity 99.8%; Pred. No. 2e-229;
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 MAFARLLRGLSPGLLGRGVCAGAMAPPCRFVLELPDCTLAHFALGADAPGDADAPDP 60
Db      1 MAFARLLRGLSPGLLGRGVCAGAMAPPCRFVLELPDCTLAHFALGADAPGDADAPDP 60
QY      61 RLAAALGPPERSYSLCVPVTPDAGCGARVRAARLHQRLRHQRRGPFQRCQLRLLCYCP 120
Db      61 RLAAALGPPERSYSLCVPVTPDAGCGARVRAARLHQRLRHQRRGPFQRCQLRLLCYCP 120
QY      121 GQAGAGAOQGFLLRDPDPTROALLLELLGACQAPRHLGFEFADPRGQWQRLWEVQ 180
Db      121 GQAGAGAOQGFLLRDPDPTROALLLELLGACQAPRHLGFEFADPRGQWQRLWEVQ 180
QY      181 DGRRLQVGCAQVVPPEPPLHPVDPDLPSSVVPDREAAARVLEECTSFIPARAVLDLV 240
Db      181 DGRRLQVGCAQVVPPEPPLHPVDPDLPSSVVPDREAAARVLEECTSFIPARAVLDLV 240
QY      241 DQCPKQIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCICQWRKIFDEPT 300
Db      241 DQCPKQIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCICQWRKIFDEPT 300
QY      301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATVAIATEVSGGLQHLPPAHPVY 360
Db      301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATVAIATEVSGGLQHLPPAHPVY 360
QY      361 QWPEDLKPDLLILLTVSPEERLQRLQGRGMKTRREEALEANSVFRQKVMESYQRMENP 420
Db      361 QWPEDLKPDLLILLTVSPEERLQRLQGRGMKTRREEALEANSVFRQKVMESYQRMENP 420
QY      421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
Db      421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449

RESULT 5
ADI40851
ID      ADI40851 standard; protein; 449 AA.
XX
AC      ADI40851;
XX
DT      22-APR-2004 (first entry)
XX
DE      Human kinase and phosphatase KPP-17 protein SEQ ID NO:17.
XX
KW      human; kinase and phosphatase; KPP; enzyme; cardiovascular;
KW      antiarteriosclerotic; hypotensive; vasotropic; antiinflammatory;
KW      antidiabetic; antiallergic; antidiabetic; nephrotropic; antitumor;
KW      antithyroid; dermatological; antidiabetic; osteopathic; uropathic;
KW      gastrointestinal; neuroprotective; antiparkinsonian; neurotropic;
KW      ophthalmological; antirheumatic; antiparasympathetic; haemostatic;
KW      anticonvulsant; hepatotropic; CNS; antiparasitic; antihelminthic; antibacterial;
KW      cytostatic; antitubercular; antiparasitic; antihelminthic; antibacterial;
KW      virucide; protozoacide; fungicide; gene therapy; kinase modulator;
KW      phosphatase modulator; cardiovascular disease; immune system disorder;
KW      neurological disorder; growth and development disorder;
KW      cell proliferative disorder; infection.
XX
OS      Homo sapiens.
XX
PN      WO2004009778-A2.
XX
PD      29-JAN-2004.
XX
PF      18-JUL-2003; 2003WO-US022650.
XX
PR      19-JUL-2002; 2002US-0397354P.
PR      02-AUG-2002; 2002US-0400509P.
PR      02-AUG-2002; 2002US-0400783P.
PR      15-AUG-2002; 2002US-0404027P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Hafalia AJA, Emerling BM, Kable AE, Richardson TW, Becha SD;

```


PI Baughn MR, Tang YT, Lal PG, Lee SY, Griffin JA, Khare R;
PI Elliott VS, Jin P, Hawkins PR, Swarnakar A, Chawla NK, Tran UK;
PI Gururajan R, Ding L, Marquis JP, Thornton MB, Forsythe JJ, Lee EA;
PI Gietzen KJ, Ramkumar J;
XX
XX
XX WPI; 2004-132950/13.
DR N-PSDB; ADI40904.
XX
XX
XX New human kinases and phosphatases, useful for diagnosing, treating or
PT preventing atherosclerosis, hypertension, AIDS, allergy, multiple
PT sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
PT or hepatitis.
XX
XX
XX Claim 1; SEQ ID NO 17; 330pp; English.
XX
XX
XX The present sequence represents a human kinase and phosphatase (KPP)
CC protein. KPP sequences have cardiovascular, antiarteriosclerotic,
CC hypotensive, vasotropic, antiinflammatory, antiangular, anti-HIV,
CC antiallergic, antiasthmatic, immunosuppressive, antithyroid,
CC dermatological, antidiabetic, nephrotropic, antigout, gastrointestinal,
CC neuroprotective, osteopathic, antiarthritic, uropathic, ophthalmological,
CC antirheumatic, antiparkinsonian, nootropic, anticonvulsant, hepatotropic,
CC CNS, antipsoriatic, haemostatic, cytostatic, antilipemic, antiparasitic,
CC antihelminthic, antibacterial, virucide, protozoacide and fungicide
CC activities, and can be used in gene therapy, and as kinase modulators and
CC phosphatase modulators. KPP proteins, polynucleotides, agonists and
CC antagonists can be used for diagnosing, treating or preventing disorders
CC associated with aberrant expression of KPP, such as cardiovascular
CC diseases (e.g. atherosclerosis, hypertension, vasculitis, angina pectoris
CC or congestive heart failure), immune system disorders (e.g. AIDS,
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC gout, irritable bowel syndrome, multiple sclerosis, osteoarthritis,
CC osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid arthritis,
CC Sjogren's syndrome or uveitis), neurological disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, dementia or
CC epilepsy), disorders affecting growth and development (e.g. cirrhosis,
CC hepatitis, mixed connective tissue disease, psoriasis or primary
CC thrombocytopenia), cell proliferative disorders (e.g.
CC hypercholesterolaemia, hyperlipidaemia or cancer), or viral, bacterial,
CC fungal, parasitic, protozoan or helminthic infections. The KPP and
CC polynucleotides are also useful in assessing the effects of exogenous
CC compounds on the expression of nucleic acids and kinases and
CC phosphatases, or for facilitating the drug discovery process, including
CC determination of efficacy, dosage, toxicity and pharmacology. The
CC polynucleotides encoding KPP are useful for creating transgenic animals
CC to model human disease.
XX
XX
XX Sequence 449 AA;
Query Match 99.5%; Score 2350; DB 8; Length 449;
Best Local Similarity 99.8%; Pred. No. 2e-229;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAFARLLRGPISGPIILGRGVCAGAMAPPCFVLELPDCTLAHFAFGADADGADADPP 60
Db 1 MAFARLLRGPISGPIILGRGVCAGAMAPPCFVLELPDCTLAHFAFGADADGADADPP 60
QY 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRGPFQRCQLRLCYCP 120
Db 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRGPFQRCQLRLCYCP 120
QY 121 GGAGGAQCGFLRLDPDPTDTRQALLEGACQAPRPHLGEFEADPRGQLWQRLWEVQ 180
Db 121 GGAGGAQCGFLRLDPDPTDTRQALLEGACQAPRPHLGEFEADPRGQLWQRLWEVQ 180
QY 181 DGRRLVGCAQVVPPEPLHPVVDLPSSVVPDREARAVLEECTSTPEARAVLDLV 240
Db 181 DGRRLVGCAQVVPPEPLHPVVDLPSSVVPDREARAVLEECTSTPEARAVLDLV 240
QY 241 DQCPKQIQKGFQVWAIEGLDGTGKTTVTSVADSLKAVLLKPPSCIGQWRKIFDDEPT 300
Db 241 DQCPKQIQKGFQVWAIEGLDGTGKTTVTSVADSLKAVLLKPPSCIGQWRKIFDDEPT 300

QY 301 IIRAFYSLGNYIVASEIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHHPVY 360
Db 301 IIRAFYSLGNYIVASEIAKESAKSPVIVDRYWHSTATYATATVSGGLQHLPPAHHPVY 360
QY 361 QWPEDLKPDILILLITVSPERLQRLQGRGMEKTRERAELEANSVFRKVMYSQRMENP 420
Db 361 QWPEDLKPDILILLITVSPERLQRLQGRGMEKTRERAELEANSVFRKVMYSQRMENP 420
QY 421 GCHVVDAVSPREKVLQTVLSLIQNSFSEP 449
Db 421 GCHVVDAVSPREKVLQTVLSLIQNSFSEP 449

RESULT 6

ADP84606
ID ADP84606 standard; protein; 449 AA.

AC ADP84606;
XX
XX 09-SEP-2004 (first entry)
XX Human breast-specific protein #110.
XX human; breast-specific protein; breast cancer.
XX Homo sapiens.
XX WO2004053077-A2.
XX 24-JUN-2004.
XX 05-DEC-2003; 2003WO-US038815.
XX 05-DEC-2002; 2002US-0431123P.
XX (DIAD-) DIADEXUS INC.

PI Macina RA, Turner LR, Sun Y, Chen H, Rodriguez M;
DR WPI; 2004-468848/44.
DR N-PSDB; ADP84474.

XX New breast specific nucleic acid molecules and polypeptides useful for
PT diagnosing, preventing or treating breast cancer, for producing
PT transgenic animals or cells, or for research purposes.

XX Claim 12; SEQ ID NO 205; 521pp; English.

XX The invention comprises the amino acid and coding sequences of human
CC breast-specific proteins. The DNA and protein sequences of the invention
CC are useful for the diagnosis, treatment and prevention of breast cancer.
CC The present amino acid sequence represents a human breast-specific
CC protein of the invention.

XX Sequence 449 AA;

Query Match 99.5%; Score 2350; DB 8; Length 449;
Best Local Similarity 99.8%; Pred. No. 2e-229;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAFARLLRGPISGPIILGRGVCAGAMAPPCFVLELPDCTLAHFAFGADADGADADPP 60
Db 1 MAFARLLRGPISGPIILGRGVCAGAMAPPCFVLELPDCTLAHFAFGADADGADADPP 60
QY 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRGPFQRCQLRLCYCP 120
Db 61 RLAAALGPERSYSLCVPTPDAGCGARVRAARLHQLRRLHQLRRGPFQRCQLRLCYCP 120
QY 121 GGAGGAQCGFLRLDPDPTDTRQALLEGACQAPRPHLGEFEADPRGQLWQRLWEVQ 180
Db 121 GGAGGAQCGFLRLDPDPTDTRQALLEGACQAPRPHLGEFEADPRGQLWQRLWEVQ 180

QY 181 DGRRLQVGCAQVVPPEPLHFWPDLPSVVPFPPREARAVLEECTSTFPEARVLDLV 240
 DB |||||
 DB 181 DGRRLQVGCAQVVPPEPLHFWPDLPSVVPFPPREARAVLEECTSTFPEARVLDLV 240
 QY 241 DQCPKIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPT 300
 DB |||||
 DB 241 DQCPKIQKGFQVVAIEGLDGTGKTTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPT 300
 QY 301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATYATIEVSGGLQHLPPAHPVY 360
 DB |||||
 DB 301 IIRAFYSLGNIVASEIAKESAKSPVIVDRYWHSTATYATIEVSGGLQHLPPAHPVY 360
 QY 361 QWPEDLKPDLLILLTVSPEERLORLQGRGMKTRAELEANSVPRQVMSYQRMENP 420
 DB |||||
 DB 361 QWPEDLKPDLLILLTVSPEERLORLQGRGMKTRAELEANSVPRQVMSYQRMENP 420
 QY 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
 DB |||||
 DB 421 GCHVVDASPSREKVLQTVLSLQNSFSEP 449
 RESULT 7
 ADC39136
 ID ADC39136 standard; protein; 422 AA.
 XX ADC39136;
 AC
 DT 18-DEC-2003 (first entry)
 XX
 DE Novel human NOVX polypeptide SEQ ID NO: 78.
 XX
 XX antidiabetic; cytostatic; immunomodulator; anorectic; antilipemic;
 KW nootropic; neuroprotective; immunostimulant; antiparkinsonian; anti-HIV;
 KW antiasthmatic; antiinflammatory; hypotensive; antiarteriosclerotic;
 KW hemostatic; osteopathic; gene therapy; NOVX; diabetes; obesity; cancer;
 KW lymphoma; uterus cancer; prostate cancer; dyslipidemia; anorexia;
 KW wasting disorder; Alzheimer's disease; Parkinson's disorder; cachexia;
 KW cardiomyopathy; AIDS; asthma; Crohn's disease; multiple sclerosis;
 KW hypertension; atherosclerosis; hemophallia; graft-versus-host disease;
 KW Albright hereditary osteodysplasy.
 XX
 OS Homo sapiens.
 XX
 XX WO2003010327-A2.
 XX
 XX 06-FEB-2003.
 XX
 XX 02-MAY-2002; 2002WO-US014199.
 XX
 XX 02-MAY-2001; 2001US-0288063P.
 XX
 XX 03-MAY-2001; 2001US-0288395P.
 XX
 XX 07-MAY-2001; 2001US-0289087P.
 XX
 XX 09-MAY-2001; 2001US-0289817P.
 XX
 XX 09-MAY-2001; 2001US-0289818P.
 XX
 XX 11-MAY-2001; 2001US-0290194P.
 XX
 XX 14-MAY-2001; 2001US-0290753P.
 XX
 XX 15-MAY-2001; 2001US-0291181P.
 XX
 XX 16-MAY-2001; 2001US-0291243P.
 XX
 XX 18-MAY-2001; 2001US-0292001P.
 XX
 XX 21-MAY-2001; 2001US-0292374P.
 XX
 XX 22-MAY-2001; 2001US-0292587P.
 XX
 XX 23-MAY-2001; 2001US-0293107P.
 XX
 XX 23-MAY-2001; 2001US-0293747P.
 XX
 XX 23-MAY-2001; 2001US-0294109P.
 XX
 XX 29-MAY-2001; 2001US-0294110P.
 XX
 XX 30-MAY-2001; 2001US-0294434P.
 XX
 XX 31-MAY-2001; 2001US-0294827P.
 XX
 XX 12-JUL-2001; 2001US-0304879P.
 XX
 XX 31-JUL-2001; 2001US-0308901P.
 XX
 XX 14-AUG-2001; 2001US-031270P.
 XX
 XX 17-AUG-2001; 2001US-0313416P.
 XX
 XX 10-SEP-2001; 2001US-0318463P.
 XX
 XX 27-SEP-2001; 2001US-0325683P.
 XX

PR 18-OCT-2001; 2001US-0330292P.
 PR 28-NOV-2001; 2001US-033873P.
 PR 03-DEC-2001; 2001US-0336909P.
 PR 03-DEC-2001; 2001US-0337552P.
 PR 21-FEB-2002; 2002US-0359245P.
 PR 01-MAY-2002; 2002US-00136826.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Miller CE, Kekuda R, Malyankar UM, Li L, Pena CEA, Spytek KA;
 PI Gorman L, Guo X, Fernandes ER, Smithson G, Stone DJ, Zerhusen BD;
 PI Patturajan M, Anderson DW, Mezes PS, Feyman JA, Macdougall JR;
 PI Padigar M, Rastelli L, Shenoy SG, Gerlach VL, Shimkets RA, Zhong M;
 PI Edinger SR, Elleman K;
 XX
 DR WPI; 2003-239445/23.
 DR N-P8DB; ADC39135.
 XX
 XX New NOVX polypeptides and polynucleotides, useful in gene therapy,
 PT particularly for treating or preventing a syndrome associated with a
 PT human disease e.g. diabetes, obesity, cancer, Alzheimer's disease,
 PT hypertension or hemophilia.
 XX
 PS Claim 1; SEQ ID NO 78; 748pp; English.
 XX
 XX The invention relates to new isolated NOVX polypeptides, the genes
 CC encoding them or sequences having at least 95% identity to the amino acid
 CC or nucleotide sequences. The NOVX polypeptide is useful as a therapeutic,
 CC particularly in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, which includes a pathology associated
 CC with NOVX polypeptide. The NOVX polypeptide is particularly useful for
 CC treating, preventing or alleviating pathology associated with NOVX
 CC polypeptide in a mammal, e.g. a human. The NOVX nucleic acid and
 CC polypeptide are especially useful for treating or preventing e.g.
 CC diabetes, obesity, cancers (e.g. lymphoma, uterus cancer or prostate
 CC cancer), dyslipidemias, anorexia, wasting disorders, Alzheimer's disease,
 CC Parkinson's disorder, cachexia, cardiomyopathy, AIDS, asthma, Crohn's
 CC disease, multiple sclerosis, hypertension, atherosclerosis, hemophilia,
 CC graft-versus-host disease or Albright hereditary osteodysplasy. The DNA
 CC encoding the protein is useful in gene therapy for treating the above
 CC conditions. These are also useful in developing powerful assay system for
 CC functional analysis of various human disorders, as well as in diagnostic
 CC applications. This sequence represents one of the NOVX proteins of the
 CC invention.
 XX
 SQ Sequence 422 AA;
 Query Match 92.2%; Score 2177.5; DB 7; Length 422;
 Best Local Similarity 99.1%; Pred. No. 6.1e-212;
 Matches 420; Conservative 0; Mismatches 1; Indels 3; Gaps 2;
 QY 27 MAPPCRFVLEPDCCTLAHFALGADADPGDADDPRLAALLGPPERSYSLCVPTPDAGCG 86
 DB 1 MAPPCRFVLEPDCCTLAHFALGADADPGDADDPRLAALLGPPERSYSLCVPTPDAGCG 60
 QY 87 ARVRAARLHORLHQRGPFQRCQLRLCYCPGGAAGAGCGFLRDLDDPTDQAL 146
 DB 61 ARVRAARLHORLHQRGPFQRCQLRLCYCPGGAAGAGCGFLRDLDDPTDQAL 120
 QY 147 LELLGACQAPRPHLGEFEADPRGQLWQRLEWVQDGRRLQVGCQVVPVPEPLHPVVD 206
 DB 121 LELLGACQAPRPHLGEFEADPRGQLWQRLEWVQDGRRLQVGCQVVPVPEPLHPVVD 180
 QY 207 LPSSVVPDREARAVLEECTSTFPEARAVLDLDVDCPKIOKGFQVVAIEGLDAT-GK 265
 DB 181 LPSSVVPDREARAVLEECTSTFPEARAVLDLDVDCPKIOKGFQVVAIEGLDATGK 240
 QY 266 TTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPTIIRAFYSLGNIVASEIAKESAKS 325
 DB 241 TTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPTIIRAFYSLGNIVASEIAKESAKS 300
 QY 326 FVIVDSYWHSTATYATIEVSGGLQHLPPAHPVYQWPEDLKPDLLILLTVSPEERLOR 385

Db 301 PVIDR-HSTATYATATEVSGGLQHLPAHPHYQWPDLLKLPDILLITVSPERLQ 358
 QY 386 LQGRGMEKTRAEAELEANSVFRQKVMYSORMENPCGHVVDAAPSREKYLQTVLSIQNS 445
 Db 359 LQGRGMEKTRAEAELEANSVFRQKVMYSORMENPCGHVVDAAPSREKYLQTVLSIQNS 418
 QY 446 FSEP 449
 Db 419 FSEP 422

RESULT 8
 ADH42559
 ID ADH42559 standard; protein; 422 AA.
 AC ADH42559;
 XX
 XX
 DT 25-MAR-2004 (first entry)
 XX
 DE Novel human protein NOV78a.
 XX
 KW cardiovascular; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
 KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;
 KW antiparkinsonian; antiaschmatic; antiinfertility; cardiomyopathy;
 KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;
 KW multiple sclerosis; graft-versus-host disease; Alzheimer's disease;
 KW Parkinson's disease; asthma; fertility disorder; chromosome mapping;
 KW tissue typing; preventive medicine; pharmacogenomic; vaccine.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO2003102159-A2.
 XX
 PD
 PF 11-DEC-2003.
 XX
 PF 04-JUN-2003; 2003WO-US017573.
 XX
 PF 04-JUN-2002; 2002US-0385490P.
 PR 04-JUN-2002; 2002US-0385615P.
 PR 04-JUN-2002; 2002US-0385755P.
 PR 05-JUN-2002; 2002US-0386041P.
 PR 06-JUN-2002; 2002US-0386355P.
 PR 06-JUN-2002; 2002US-0386357P.
 PR 06-JUN-2002; 2002US-0386447P.
 PR 06-JUN-2002; 2002US-0386459P.
 PR 06-JUN-2002; 2002US-0386465P.
 PR 06-JUN-2002; 2002US-0386701P.
 PR 07-JUN-2002; 2002US-0386796P.
 PR 07-JUN-2002; 2002US-0386931P.
 PR 07-JUN-2002; 2002US-0387078P.
 PR 07-JUN-2002; 2002US-0387081P.
 PR 07-JUN-2002; 2002US-0387083P.
 PR 10-JUN-2002; 2002US-0387429P.
 PR 10-JUN-2002; 2002US-0387540P.
 PR 10-JUN-2002; 2002US-0387866P.
 PR 11-JUN-2002; 2002US-0387866P.
 PR 11-JUN-2002; 2002US-0387610P.
 PR 11-JUN-2002; 2002US-0387659P.
 PR 11-JUN-2002; 2002US-0387668P.
 PR 11-JUN-2002; 2002US-0387668P.
 PR 11-JUN-2002; 2002US-0387668P.
 PR 11-JUN-2002; 2002US-0387859P.
 PR 12-JUN-2002; 2002US-0387934P.
 PR 12-JUN-2002; 2002US-0387960P.
 PR 12-JUN-2002; 2002US-0388022P.
 PR 12-JUN-2002; 2002US-0388096P.
 PR 12-JUN-2002; 2002US-0388432P.
 PR 12-JUN-2002; 2002US-0388479P.
 PR 13-JUN-2002; 2002US-0389123P.
 PR 14-JUN-2002; 2002US-0389120P.
 PR 14-JUN-2002; 2002US-0389146P.
 PR 17-JUN-2002; 2002US-0389742P.
 PR 18-JUN-2002; 2002US-0389604P.

PR 18-JUN-2002; 2002US-0389884P.
 PR 19-JUN-2002; 2002US-0390006P.
 PR 19-JUN-2002; 2002US-0390144P.
 PR 19-JUN-2002; 2002US-0390209P.
 PR 25-JUN-2002; 2002US-0391726P.
 PR 06-AUG-2002; 2002US-0401628P.
 PR 09-AUG-2002; 2002US-0402268P.
 PR 12-AUG-2002; 2002US-0402822P.
 PR 13-AUG-2002; 2002US-0403458P.
 PR 15-AUG-2002; 2002US-0403617P.
 PR 15-AUG-2002; 2002US-0403732P.
 PR 26-AUG-2002; 2002US-0406182P.
 PR 12-SEP-2002; 2002US-0410085P.
 PR 13-SEP-2002; 2002US-0410505P.
 PR 23-SEP-2002; 2002US-0412955P.
 PR 30-SEP-2002; 2002US-0415195P.
 PR 23-OCT-2002; 2002US-0420627P.
 PR 23-OCT-2002; 2002US-0420718P.
 PR 24-OCT-2002; 2002US-0420852P.
 PR 31-OCT-2002; 2002US-0422750P.
 PR 01-NOV-2002; 2002US-0423095P.
 PR 05-NOV-2002; 2002US-0423748P.
 (CURA-) CURAGEN CORP.
 PA
 XX
 XX
 PI Alsbrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;
 PI Burgess CE, Casman SJ, Catterton E, Dhanabal M, Edinger SR;
 PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
 PI Grossie NW, Gunther E, Guo X, Gusev VY, Herrmann JL, Ji W, Kekuda R;
 PI Khrantsov NV, Larochelle WJ, Li L, Liang H, Low K, Macdougall JR;
 PI MacLachlan T, Malyankar UM, McQueeney K, Mezick AJ, Miller CE;
 PI Millet I, Padigaru M, Patturajan M, Peyman JA, Qian X, Rastelli L;
 PI Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;
 PI Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;
 PI Wolenc AR, Zhong M, Zhong H;
 XX
 XX
 DR WPI: 2004-053467/05.
 DR N-PSDB; ADH42558.
 XX
 XX
 PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
 treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
 PT pharmacogenomics.
 XX
 PS Claim 2; SEQ ID NO 1112; 1503pp; English.
 XX
 CC The invention relates to 566 new isolated human polypeptides and their
 CC encoding genes, sequences that are at least 95% identical to these or
 CC sequences comprising one or more conservative substitutions in these. The
 CC polypeptide, polynucleotide and antibodies against the polypeptides are
 CC useful in diagnosing, treating or preventing NOVX-associated disorders,
 CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
 CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
 CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
 CC The nucleic acids are further used as hybridization probes, in chromosome
 CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
 CC polypeptides are also useful as vaccines. This sequence represents an
 CC example of the polypeptide of the invention.
 XX
 SQ Sequence 422 AA;

Query Match 92.2%; Score 2177.5; DB 8; Length 422;
 Best Local Similarity 99.1%; Pred. No. 6.1e-212;
 Matches 420; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 27 MAPPCRFVLELPDCTLAHFALGADAPGDADAPDLAALGPPERSYSLCVPVTPDAGCG 86
 Db 1 MAPPCRFVLELPDCTLAHFALGADAPGDADAPDLAALGPPERSYSLCVPVTPDAGCG 60
 QY 87 ARVRAARLHQRLLHQLRRGPFQRCOLLRLCYCPGQAGGAGQGLLRPLDLPDPTQAL 146
 Db 61 ARVRAARLHQRLLHQLRRGPFQRCOLLRLCYCPGQAGGAGQGLLRPLDLPDPTQAL 120

QY 147 LELLGACQAPRPHLGEFEADPRGOLWQRLWEVDGRRLOVGAQVVPPEPLHPVDP 206
 DB 121 LELLGACQAPRPHLGEFEADPRGOLWQRLWEVDGRRLOVGAQVVPPEPLHPVDP 180
 QY 207 LPSSVFPDREARAVLEECTSFIPPEARAVLDLDVDCPKQIQKGKQVVAIEGLDAT-GK 265
 DB 181 LPSSVFPDREARAVLEECTSFIPPEARAVLDLDVDCPKQIQKGKQVVAIEGLDATGK 240
 QY 266 TTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFYSLGNYIVASEIAKESAKS 325
 DB 241 TTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFYSLGNYIVASEIAKESAKS 300
 QY 326 PVIIVRYWSTATYATATEVSGLOHLPAAHHPVYQWPEDLKPDILLITVSPERLOR 385
 DB 301 PVIIVDR--HSTATYATATEVSGLOHLPAAHHPVYQWPEDLKPDILLITVSPERLOR 358
 QY 386 LQGRGMEKTRAEAELEANSVFRQKVMSYQRMENPGCHVVVDASPSREKVLQTVLSIQNS 445
 DB 359 LQGRGMEKTRAEAELEANSVFRQKVMSYQRMENPGCHVVVDASPSREKVLQTVLSIQNS 418
 QY 446 FSEP 449
 DB 419 FSEP 422

RESULT 9
 AA38614
 ID AA38614 standard; protein; 395 AA.

AC AA38614;

DT 04-DEC-2003 (first entry)

XX Mouse thymidylate kinase protein.

XX Mouse; thymidylate kinase; EC 2.7.4.9; ATP:dtMP phosphotransferase; TK;
 KW enzyme; cancer; gene therapy.

XX Mus musculus.

XX WO2003064642-A1.

XX 07-AUG-2003.

XX 27-JAN-2003; 2003WO-EP000783.

XX 28-JAN-2002; 2002US-0351424P.

XX 25-NOV-2002; 2002US-0428711P.

XX (FARB) BAYER AG.

XX Liou J;

XX WPI; 2003-646151/61.

XX New polynucleotide and its encoded thymidylate kinase (TK), useful for
 PT identifying modulators of TK activity (e.g. agonists or antagonists) that
 PT provides therapeutic effects, and in gene therapy for treating cancers.

XX Disclosure; Page 88-89; 110pp; English.

XX The invention relates to human thymidylate kinase (TK) polypeptides (EC
 CC 2.7.4.9) and polynucleotides. TK is also referred to as ATP:dtMP
 CC phosphotransferase. TK sequences are useful for identifying test
 CC compounds, that may act as agonists or antagonists at the receptor site
 CC and which can be regulated to provide therapeutic effects. Vectors
 CC comprising TK polynucleotides are useful in the preparation of a
 CC medicament for modulating the activity of TK in a disease, particularly
 CC cancer. The invention is also useful in gene therapy. The present
 CC sequence is mouse thymidylate kinase protein. This sequence is used in
 CC the invention

XX Sequence 395 AA;

Query Match 65.7%; Score 1552; DB 7; Length 395;
 Best Local Similarity 77.8%; Pred. No. 2.2e-148;
 Matches 294; Conservative 36; Mismatches 48; Indels 0; Gaps 0;

QY 68 PERSYSLCVPTPDAGCGARVARAAARLHORLHQLRRGPFQRCQLRLLCYCPGGQAGGA 127
 DB 16 PQGAATRLCVPLAPGCGCGPRVQAARVHRLQLQRRGFLQRCQLSKLLGYGPGDQAGEA 75
 QY 128 QQGFLLRDLDDPTDTRQALLELLGACQAPRPHLGEFEADPRGOLWQRLWEVDGRRLOV 187
 DB 76 QHGFLLRDLDDPTDTRDILLQLGSCQEAARFQAEFFQADSQGLLWQRLWEVDGRRLOV 135
 QY 188 GCAQVVPPEPLHPVVPDLPSSVFPDREARAVLEECTSFIPPEARAVLDLDVDCPKQI 247
 DB 136 DCACVLPAGEPHLHLLPDLNLNSAVFQDRDARAVLEECTSFIPPEARAVLDLDVDCPKV 195
 QY 248 QKGKQVVAIEGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFY 307
 DB 196 QKGKQVVAIEGLDATGKTTLTQSVSESLKAVLLQSPPCISQWRKIFDDEPTIIRAFY 255
 QY 308 SLGNYIVASEIAKESAKSPVIIVDRYWHSTATYATATEVSGLOHLPAAHHPVYQWPEDL 367
 DB 256 SLGNYIVASEIAKESNTFFVIVDRYWHSTATYATATEVSGLOHLPAAHHPVYQWPEDL 315
 QY 368 KPDILLITVSPERLORLQGRGMEKTRAEAELEANSVFRQKVMSYQRMENPGCHVVDA 427
 DB 316 KPDILLITVSPERLORLQGRGMEKTRAEAELEANSVFRQKVMSYQRMENPGCHVVDA 375
 QY 428 SPSREKVLQTVLSIQNS 445
 DB 376 SPSRETVLQKVLQIQQS 393

RESULT 10

ADJ76250

ID ADJ76250 standard; protein; 395 AA.

AC ADJ76250;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:1502.

XX bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker.

XX Mus musculus.

XX EP1394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.

XX Claim 16; SEQ ID NO 1502; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma

or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

XX Sequence 395 AA;

Query Match 65.7%; Score 1552; DB 8; Length 395;
 Best Local Similarity 77.8%; Pred. No. 2.2e-148;
 Matches 294; Conservative 36; Mismatches 48; Indels 0; Gaps 0;

QY 68 PPERSYSLCVPTDAGCGARVRAARLHQLRRGPFQRCOLLRLCYCPGGAGGA 127
 Db 16 PQGATRLCVLAPGEGCGPRVQAARVHRLQLQRRGLQRCQLKLYGEGDQAGEA 75

QY 128 QGFLRLDPLDPTDTRQALLLELGGACQAPRHLGFEADPRGQLWRLWEVQDGRRLQV 187
 Db 76 QHGFLLRDCDHPDTRDLRLQLLGGSCQEAARPCQAEFQADSQGLLWQLWELQGRQVQV 135

QY 188 GCAQVVPVPEPLHPVVPDLPSVVPDREARAVLEECTSFTEPEARVLDLVDQCPKQI 247
 Db 136 DCACVLPAQEPHLHPDPLLSAVFQDRDARAVLEECTSFTEPEARVLDLVDQCPKSV 195

QY 248 QKGKFOVAIEGLDAGTKTQVTSVADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFY 307
 Db 196 QKGKFOVAIEGLDAGTKTTLTQSVESLKAVLLQSPFPCISQWRKIFDDEPTIIRAFY 255

QY 308 SLGNIVASEIAKESAKSPVIVDRYWHSTATYAIAATEVSGGLQHLPPAHPVYQWPEDL 367
 Db 256 SLGNVLVASEIAKESSTNFFVIVDRYWHSTATYAIAATEVSGGLQYLPAPHPVYQWPGD 315

QY 368 KPDILLITVSPERLQRCGRMEKTRAEAELEANSVFRQKVMYSIQRMENPGCHVDA 427
 Db 316 KPDVLLITVNSERVRLQGRGQEKTKBEAELEANNVFRQKVMYSIQRMENPNSCHLVD 375

QY 428 SPSREKVLQTVLSLIONS 445
 Db 376 SPSRETVLQVLELIQSS 393

RESULT 11
 ADC39138
 ID ADC39138 standard; protein; 251 AA.
 XX ADC39138;
 AC ADC39138;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 XX Novel human NOVX polypeptide SEQ ID NO: 80.

XX antidiabetic; cytostatic; immunomodulator; anorectic; antilipemic;
 KW neurotropic; neuroprotective; immunostimulant; antiparkinsonian; anti-HIV;
 KW antiasthmatic; antiinflammatory; hypotensive; antiarteriosclerotic;
 KW hemostatic; osteopathic; gene therapy; NOVX; diabetes; obesity; cancer;
 KW lymphoma; uterine cancer; prostate cancer; dyslipidemia; anorexia;
 KW wasting disorder; Alzheimer's disease; Parkinson's disorder; cachexia;
 KW cardiomyopathy; AIDS; asthma; Crohn's disease; multiple sclerosis;
 KW hypertension; atherosclerosis; hemophilia; graft-versus-host disease;
 KW Albright hereditary osteodystrophy.

XX Homo sapiens.

XX WO2003010327-A2.

XX 06-FEB-2003.

XX 02-MAY-2002; 2002WO-US014199.

XX 02-MAY-2001; 2001US-0288063P.

XX 03-MAY-2001; 2001US-0288395P.

XX 07-MAY-2001; 2001US-0289087P.

XX 09-MAY-2001; 2001US-0289817P.

XX 11-MAY-2001; 2001US-0289818P.

XX 14-MAY-2001; 2001US-0290194P.

XX 15-MAY-2001; 2001US-0291181P.

XX 16-MAY-2001; 2001US-0291243P.

XX 18-MAY-2001; 2001US-0292001P.

XX 21-MAY-2001; 2001US-0292374P.

XX 22-MAY-2001; 2001US-0292587P.

XX 23-MAY-2001; 2001US-0293107P.

XX 25-MAY-2001; 2001US-0293747P.

XX 29-MAY-2001; 2001US-0294109P.

XX 30-MAY-2001; 2001US-0294110P.

XX 31-MAY-2001; 2001US-0294434P.

XX 12-JUL-2001; 2001US-0294827P.

XX 31-JUL-2001; 2001US-0304879P.

XX 14-AUG-2001; 2001US-030901P.

XX 17-AUG-2001; 2001US-0312270P.

XX 10-SEP-2001; 2001US-0313416P.

XX 27-SEP-2001; 2001US-0318463P.

XX 18-OCT-2001; 2001US-0325683P.

XX 28-NOV-2001; 2001US-0330292P.

XX 03-DEC-2001; 2001US-0333873P.

XX 21-FEB-2002; 2002US-0336909P.

XX 01-MAY-2002; 2002US-0359245P.

XX 01-MAY-2002; 2002US-00136826.

(CURA-) CURAGEN CORP.

XX Miller CE, Kekuda R, Malyankar UM, Li L, Pena CEA, Spytek KA;
 PI Gorman L, Guo X, Fernandes ER, Smithson G, Stone DU, Zerhusen BD;
 PI Patturajan M, Anderson DW, Mezes PS, Peyman JA, Macdougall JR;
 PI Padgaru M, Rastelli L, Shenoy SG, Gerlach VL, Shimkets RA, Zhong M;
 PI Edinger SR, Eilerman K;
 XX WPI; 2003-239445/23.

XX N-PSDB; ADC39137.

XX New NOVX polypeptides and polynucleotides, useful in gene therapy,
 PT particularly for treating or preventing a syndrome associated with a
 PT human disease e.g. diabetes, obesity, cancer, Alzheimer's disease,
 PT hypertension or hemophilia.

XX Claim 1; SEQ ID NO 80; 748pp; English.

XX The invention relates to new isolated NOVX polypeptides, the genes
 CC encoding them or sequences having at least 95% identity to the amino acid
 CC or nucleotide sequences. The NOVX polypeptide is useful as a therapeutic,
 CC particularly in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, which includes a pathology associated

PI Maclachlan T, Malyankar UM, McQueeney K, Mezick AJ, Miller CE;
 PI Millet I, Padigaru M, Patturajan M, Peyman JA, Qian X, Rastelli L;
 PI Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;
 PI Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;
 PI Wolenc AR, Zhong M, Zhong H;
 XX WPI: 2004-053467/05.
 DR N-PSDB; ADH42564.
 XX New NOVX polypeptides and nucleic acid molecules useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
 PT pharmacogenomics.
 XX
 PS Claim 2; SEQ ID NO 118; 1503pp; English.
 XX
 CC The invention relates to 566 new isolated human polypeptides and their
 CC encoding genes, sequences that are at least 95% identical to these or
 CC sequences comprising one or more conservative substitutions in these. The
 CC polypeptide, polynucleotide and antibodies against the polypeptides are
 CC useful in diagnosing, treating or preventing NOVX-associated disorders,
 CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
 CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
 CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
 CC The nucleic acids are further used as hybridization probes, in chromosome
 CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
 CC polypeptides are also useful as vaccines. This sequence represents an
 CC example of the polypeptide of the invention.
 XX
 SQ Sequence 251 AA;

Query Match 50.2%; Score 1185; DB 8; Length 251;
 Best Local Similarity 58.9%; Pred. No. 2.2e-111;
 Matches 249; Conservative 0; Mismatches 2; Indels 172; Gaps 1;
 QY 27 MAPPCRFVLELPDCTLAHFAFGADAPGDADAPDRLAALLGPPERSYSLCVPTPDAGCG 86
 Db 1 MAPPCRFVLELPDCTLAHFAFG----- 22
 QY 87 ARVRAARLHQRLHQLRRGFQRCQLRLCYCPGGAGGAGQGLLRDPLDPPDTRQAL 146
 Db 23 ----- 22
 QY 147 LELLGACQAPRHLGFEFADPRGLWQRLWEVDGRRLQVCAQVVPVPEPLHPVDP 206
 Db 23 ----- 22
 QY 207 LPSSVVPDREARAVLECTSFIPPEARAVLDLDVQCPKQIQKGFQVVAIEGLDATGKT 266
 Db 23 -----AVLECTSFIPPEARAVLDLDVQCPKQIQKGFQVVAIEGLDATGKT 68
 QY 267 TVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTTIIRRAFYSIGNIVASEIAKESAKSP 326
 Db 69 TVTQSAADSLKAVLLKSPSCIGQWRKIFDDEPTTIIRRAFYSIGNIVASEIAKESAKSP 128
 QY 327 VIVDRYWHSTATYATATEVSGGLQHLPPAHPVYQWPEDLKPDLLILLITVSPPELQRL 386
 Db 129 VIVDRYWHSTATYATATEVSGGLQHLPPAHPVYQWPEDLKPDLLILLITVSPPELQRL 188
 QY 387 QGRGMEKTRBEAELEANSFRQKVENSYORMENPGCHVVVDASPSREKVLQTVLSLQNSF 446
 Db 189 QGRGMEKTRBEAELEANSFRQKVENSYORMENPGCHVVVDASPSREKVLQTVLSLQNSF 248
 QY 447 SEP 449
 Db 249 SEP 251
 RESULT 13
 ID ADF76971
 XX ADF76971 standard; protein; 290 AA.
 AC ADF76971;

XX
 DT 26-FEB-2004 (first entry)
 XX Novel human secreted and transmembrane protein SeqID 646.
 DE
 XX human; PRO; membrane bound protein; membrane bound receptor;
 KW cell proliferation; cell migration; cell differentiation;
 KW mitogenic factor; survival factor; cytotoxic factor;
 KW differentiation factor; neuropeptide; hormone; cell receptor;
 KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.
 XX Homo sapiens.
 OS
 XX WO2003072035-A2.
 PN
 XX 04-SEP-2003.
 PD
 XX 21-FEB-2003; 2003WO-US005241.
 PF
 XX 22-FEB-2002; 2002US-0359461P.
 PR
 XX (GETH) GENENTECH INC.
 PA Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
 PI Williams PM, Wood WI, Wu TD;
 XX WPI: 2003-721702/68.
 DR N-PSDB; ADF76970.
 DR
 XX New PRO polypeptides, useful for diagnosing and treating an immune
 PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
 PT diabetes mellitus.
 XX
 PS Claim 10; SEQ ID NO 646; 918pp; English.
 XX
 CC This invention relates to novel nucleic acids encoding human PRO secreted
 CC and transmembrane proteins. Extracellular proteins play important roles
 CC in the formation, differentiation and maintenance of multicellular
 CC organisms. The fate of many individual cells (for example proliferation,
 CC migration or differentiation) is typically governed by information
 CC received from other cells and the immediate environment. The information
 CC is often transmitted by secreted polypeptides (for example mitogenic
 CC factors, survival factors, cytotoxic factors, differentiation factors,
 CC neuropeptides and hormones) which are received and interpreted by diverse
 CC cell receptors or membrane bound proteins. These membrane bound proteins
 CC as in the blocking of receptor-ligand interactions. The current invention
 CC provides the amino acid sequences of novel human membrane bound receptors
 CC and proteins, along with the cDNA sequences encoding them. The novel
 CC proteins of the invention may have cytostatic activities through the
 CC stimulation of chondrocytes. The nucleic acids of the invention may be
 CC useful for the manufacture of a medicament for diagnosing or treating a
 CC tumour in a mammal. In addition, they may be useful for measuring or
 CC detecting the expression of a tumour associated gene. The present
 CC sequence is the amino acid sequence of a human PRO protein of the
 CC invention.
 XX
 SQ Sequence 290 AA;

Query Match 48.1%; Score 1137; DB 7; Length 290;
 Best Local Similarity 90.9%; Pred. No. 2.1e-106;
 Matches 229; Conservative 3; Mismatches 16; Indels 4; Gaps 2;
 QY 202 PVVDPDLP-SVVFPPDREARAVLECT---SFIPPEARAVLDLDVQCPKQIQKGFQVVAI 257
 Db 39 PIWALPSPWRSKREAPRRGKSLCTVYLLYIPPEARAVLDLDVQCPKQIQKGFQVVAI 98
 QY 258 EGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTTIIRRAFYSIGNIVASE 317
 Db 99 EGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKIFDDEPTTIIRRAFYSIGNIVASE 158
 QY 318 IAKESAKSPVIVDRYWHSTATYATATEVSGGLQHLPPAHPVYQWPEDLKPDLLILLITV 377

Db 159 IAKESAKSPVVDYVHWHSTATYATATEVSGGLQHLPPAHHPVYQWPEDLKPDILLILLTV 218
QY 378 SPEERLQRLQGMKTRTEAELEANSVFRQKVENSYQRMENPGCHVVDASPSREKVLQT 437
Db 219 SPEERLQRLQGMKTRTEAELEANSVFRQKVENSYQRMENPGCHVVDASPSREKVLQT 278
QY 438 VLSLQNSFSSEP 449
Db 279 VLSLQNSFSSEP 290
RESULT 14
ADH42561
ID ADH42561 standard; protein; 200 AA.
XX
AC ADH42561;
XX
DT 25-MAR-2004 (first entry)
XX
DE Novel human protein NOV78b.
XX
KW cardiovascular; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;
KW antiparkinsonian; antischismic; antiinfertility; cardiomyopathy;
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;
KW multiple sclerosis; graft-versus-host disease; Alzheimer's disease;
KW Parkinson's disease; asthma; fertility disorder; chromosome mapping;
KW tissue typing; preventive medicine; pharmacogenomic; vaccine.
XX
OS Homo sapiens.
XX
PN WC2003102159-A2.
XX
PD 11-DEC-2003.
XX
XX
PF 04-JUN-2003; 2003WC-US017573.
XX
PR 04-JUN-2002; 2002US-0385490P.
PR 04-JUN-2002; 2002US-0385615P.
PR 04-JUN-2002; 2002US-0385755P.
PR 05-JUN-2002; 2002US-0386041P.
PR 06-JUN-2002; 2002US-0386355P.
PR 06-JUN-2002; 2002US-0386357P.
PR 06-JUN-2002; 2002US-0386447P.
PR 06-JUN-2002; 2002US-0386459P.
PR 06-JUN-2002; 2002US-0386465P.
PR 07-JUN-2002; 2002US-0386701P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0387078P.
PR 07-JUN-2002; 2002US-0387081P.
PR 10-JUN-2002; 2002US-0387083P.
PR 10-JUN-2002; 2002US-0387429P.
PR 10-JUN-2002; 2002US-0387540P.
PR 10-JUN-2002; 2002US-0387866P.
PR 11-JUN-2002; 2002US-0387606P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387659P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 12-JUN-2002; 2002US-0388432P.
PR 12-JUN-2002; 2002US-0388479P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389742P.

PR 18-JUN-2002; 2002US-0389604P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390144P.
PR 19-JUN-2002; 2002US-0390209P.
PR 25-JUN-2002; 2002US-0391726P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-040268P.
PR 12-AUG-2002; 2002US-0402822P.
PR 13-AUG-2002; 2002US-0403458P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 12-SEP-2002; 2002US-0410085P.
PR 13-SEP-2002; 2002US-0410505P.
PR 23-SEP-2002; 2002US-0412955P.
PR 30-SEP-2002; 2002US-0415195P.
PR 23-OCT-2002; 2002US-0420627P.
PR 23-OCT-2002; 2002US-0420718P.
PR 24-OCT-2002; 2002US-0420852P.
PR 31-OCT-2002; 2002US-0422750P.
PR 01-NOV-2002; 2002US-0423095P.
PR 05-NOV-2002; 2002US-0423748P.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;
PI Burgess CE, Casman SU, Catterton E, Dhanabal M, Edinger SR;
PI Ellerman K, Ethenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI Grosse WW, Gunther E, Guo X, Gusev VY, Herrmann JL, Ji W, Kekuda R;
PI Khrantsov NV, Larochele WJ, Li L, Liang H, Low K, Macdougall JR;
PI MacLachlan T, Malyankar UM, McQueeney K, Mezick AJ, Miller CE;
PI Millet I, Padigaru M, Patturajan M, Peyman JA, Qian X, Rastelli L;
PI Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;
PI Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;
PI Wolenc AR, Zhong M, Zhong H;
XX
XX WPI; 2004-053457/05.
DR N-PSDB; ADH42560.
XX
XX New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
PT pharmacogenomics.
XX
PS Claim 2; SEQ ID NO 1114; 1503pp; English.
XX
XX The invention relates to 566 new isolated human polypeptides and their
CC encoding genes, sequences that are at least 95% identical to these or
CC sequences comprising one or more conservative substitutions in these. The
CC polypeptide, polynucleotide and antibodies against the polypeptides are
CC useful in diagnosing, treating or preventing NOVX-associated disorders,
CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
CC The nucleic acids are further used as hybridization probes, in chromosome
CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
CC polypeptides are also useful as vaccines. This sequence represents an
CC example of the polypeptide of the invention.
XX
SQ Sequence 200 AA;
Query Match 41.6%; Score 982.5; DB 8; Length 200;
Best Local Similarity 99.0%; Pred. No. 6.1e-91;
Matches 193; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 256 AIEGLDAT-GKTTVTQSVADSLKAVLLKSPSCIGQWRKTFDDEPTIIRAFYSLGNIV 314
Db 3 SIEGLDATGKTTVTQSVADSLKAVLLKSPSCIGQWRKTFDDEPTIIRAFYSLGNIV 62
QY 315 ASEIAKESAKSPVVDYVHWHSTATYATATEVSGGLQHLPPAHHPVYQWPEDLKPDILL 374
Db 63 ASEIAKESAKSPVVDYVHWHSTATYATATEVSGGLQHLPPAHHPVYQWPEDLKPDILL 122

Db 123 LTVSPEERLQRLQGRGMEKTRREEALEANSVFRQKVENSYQRMENPGCHVVDASPSREKV 182
Qy 435 LQTVLSLIONSFSSEP 449
Db 183 LQTVLSLIONSFSSEP 197

Search completed: October 22, 2004, 23:30:58
Job time : 123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 10:45:02 ; Search time 7916 Seconds

(without alignments)
11835.087 Million cell updates/sec

Title: US-10-681-223-1

Perfect score: 2571

Sequence: 1 cggggcggggcgctcg.....aggaggggcttttcacccc 2571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gssi: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	885.2	34.4	1003	5	BX463706	BX463706 BX463706
2	884.2	34.4	3124	3	BC027329	BC027329 Mus muscu
3	873.2	34.0	3214	3	BC048910	BC048910 Mus muscu
4	872.4	33.9	3022	3	AK004595	AK004595 Mus muscu
5	771.4	30.0	1084	5	BX421924	BX421924 BX421924
6	771.4	30.0	1287	3	CR623116	CR623116 full-leng
7	741.8	28.9	949	6	CF272319	CF272319 AGENCOURT
8	690.2	26.8	757	4	BG501037	BG501037 602546443
9	685.2	26.7	767	4	BG497576	BG497576 601859469
10	674.8	26.2	759	4	BG497578	BG497578 602543017
11	653.2	25.4	769	5	BUI83182	BUI83182 AGENCOURT
12	646.6	25.1	705	4	BG778349	BG778349 602666287
13	646.6	25.1	833	4	B1551715	B1551715 603197411
14	621.8	24.2	844	4	BG778145	BG778145 602666482
15	613.8	23.9	720	4	BG496231	BG496231 602538182
16	607	23.6	1262	5	BQ067262	BQ067262 AGENCOURT
17	605.4	23.5	801	5	BUI60713	BUI60713 AGENCOURT
18	602.4	23.4	611	5	BU070188	BU070188 in19c07.y
19	600.4	23.4	972	4	BUI78500	BUI78500 602328334
20	594.4	23.1	596	6	CB114648	CB114648 K-EST0158
21	588.4	22.9	888	2	BF240527	BF240527 601875619
22	588.2	22.9	669	2	BE748448	BE748448 601571914
23	584.2	22.7	1333	4	BM906075	BM906075 AGENCOURT
24	583.6	22.7	765	2	BF241358	BF241358 601876319

25	569	22.1	728	4	BG779993	BG779993 602668235
26	563.6	21.9	751	2	BF102537	BF102537 601646494
27	563.6	21.9	868	2	BF243467	BF243467 601878183
28	562	21.9	700	4	B1818166	B1818166 603032614
29	553.6	21.5	615	4	BG776607	BG776607 602663751
30	549.4	21.4	616	1	AV707941	AV707941 AV707941
31	549.2	21.4	813	2	BF242831	BF242831 601878235
32	534	20.8	702	4	BG776636	BG776636 602663790
33	533.2	20.7	795	4	BG771320	BG771320 602719573
34	531.8	20.7	595	4	BG501398	BG501398 602547746
35	529.8	20.6	837	2	BF242245	BF242245 601878076
36	524.2	20.4	669	4	BG497460	BG497460 601859520
37	521.8	20.3	704	4	BG547557	BG547557 602575410
38	520.2	20.2	1016	6	BY704814	BY704814 BX704814
39	516.6	20.1	596	2	BE748211	BE748211 601571424
40	515.8	20.1	745	2	BF031764	BF031764 601558088
41	513.8	20.0	830	5	BUI66495	BUI66495 AGENCOURT
42	511.2	19.9	735	2	BF243749	BF243749 601877866
43	506	19.7	1008	4	B1918806	B1918806 603180784
44	493.8	19.2	928	4	BG180852	BG180852 602327301
45	483.8	18.8	771	4	BG255761	BG255761 602368237

ALIGNMENTS

RESULT 1
BX463706
LOCUS BX463706 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS0DN005YL22 5-PRIME, mRNA sequence.
ACCESSION BX463706
VERSION BX463706.2 GI:47071037
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1003)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT On May 22, 2003 this sequence version replaced gi:31035556.
Contract: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1573.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DN005DF11QPlsc=1573.r.

FEATURES

source
1..1003
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YL22"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 34.4%; Score 885.2; DB 5; Length 1003;
Best Local Similarity 99.0%; Pred. No. 3.2e-198;

Matches	901;	Conservative	0;	Mismatches	8;	Indels	1;	Gaps	1;
Qy	695	FTTTTGGAGAGTGTACCTCTTTTATCTCTGAAGCCCGGCGAGTGTCTGACCTGTGTCGAC	754						
Db	95	GCTTGTGCACAGTGTACCTCTTTATCTCTGAAGCCCGGCGAGTGTCTGACCTGTGTCGAC	154						
Qy	755	CAGTGCCTCCAAACAGATCCAGAAAGGAATGCCAGTTGTGCCATCGAAGACTGGAT	814						
Db	155	CAGTGCCTCCAAACAGATCCAGAAAGGAATGCCAGTTGTGCCATCGAAGACTGGAT	214						
Qy	815	GCACGGGTAAACACAGCTGACCCAGTCAGTGGCAGATTCACCTTAAGGCTGTCTCTTAA	874						
Db	215	GCACGGGTAAACACAGCTGACCCAGTCAGTGGCAGATTCACCTTAAGGCTGTCTCTTAA	274						
Qy	875	AAGTCAACCCCTCTTTCATTTGGCCAGTGGAGGAATCTTTGATGATGAACCAACTATC	934						
Db	275	AAGTCAACCCCTCTTTCATTTGGCCAGTGGAGGAATCTTTGATGATGAACCAACTATC	334						
Qy	935	ATTAGAAGAGCTTTTACTCTTTGGGCAATTATTTGTGGCCCTCCGAATAGCTTAAGAA	994						
Db	335	ATTAGAAGAGCTTTTACTCTTTGGGCAATTATTTGTGGCCCTCCGAATAGCTTAAGAA	394						
Qy	995	TCGTCCAAATCTCCTGTGATTGTAGACAGTACTGGCAGACGCGCCACTATGCCATA	1054						
Db	395	TCGTCCAAATCTCCTGTGATTGTAGACAGTACTGGCAGACGCGCCACTATGCCATA	454						
Qy	1055	GCCACTGAGTGTGGGGTCTCCAGACCTGCCCGCCAGCCCATCACCTGTGTACCAG	1114						
Db	455	GCCACTGAGTGTGGGGTCTCCAGACCTGCCCGCCAGCCCATCACCTGTGTACCAG	514						
Qy	1115	TGCCAGAGACCTGCTCAACCTGACCTTATCTGTCTGTCTACTGTGAGTCTCTGAGGAG	1174						
Db	515	TGCCAGAGACCTGCTCAACCTGACCTTATCTGTCTGTCTACTGTGAGTCTCTGAGGAG	574						
Qy	1175	AGTTTGAGAGGCTGACGGCCGGGCGATGGAGAAAGACGAGGGAAGAGCAACTTGAG	1234						
Db	575	AGTTTGAGAGGCTGACGGCCGGGCGATGGAGAAAGACGAGGGAAGAGCAACTTGAG	634						
Qy	1235	GCCAAAGTGTGTTTGTGCTAAAAGGTAGAAAATGTCTTACAGGGATGGAGATCTCTGGC	1294						
Db	635	GCCAAAGTGTGTTTGTGCTAAAAGGTAGAAAATGTCTTACAGGGATGGAGATCTCTGGC	694						
Qy	1295	TGCATGTGTGTTGTTGCTAAAAGGTAGAAAATGTCTTACAGGGATGGAGATCTCTGGC	1354						
Db	695	TGCATGTGTGTTGTTGCTAAAAGGTAGAAAATGTCTTACAGGGATGGAGATCTCTGGC	754						
Qy	1355	ATCCAGAAATAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCAGTCTAACTAGATT	1414						
Db	755	ATCCAGAAATAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCAGTCTAACTAGATT	813						
Qy	1415	AGATGTTGTTGAAACATCTACATCCACCATTTGTTATGAGTGTTCGAAATTTCTGTT	1474						
Db	814	AGATGTTGTTGAAACATCTACATCCACCATTTGTTATGAGTGTTCGAAATTTCTGTT	873						
Qy	1475	CTACAAGCTGTGTGTGGCAGAAACTGGAGACAGGCATCTTAATTTTACTTTCAGCCA	1534						
Db	874	CTACAAGCTGTGTGTGGCAGAAACTGGAGACAGGCATCTTAATTTTACTTTCAGCCA	933						
Qy	1535	TCGTACCTCTTCTGACTGATGGACCGCTCATCAAAAGGTTCCTCTCATCATGTTCCAG	1594						
Db	934	TCGTACCTCTTCTGACTGATGGACCGCTCATCAAAAGGTTCCTCTCATCATGTTCCAG	993						
Qy	1595	TGAGAGGCCA 1604							
Db	994	TGAGAGGCCA 1003							

RESULT 2
BC027329
LOCUS
DEFINITION
ACCESSION

BC027329 3124 bp mRNA linear HTC 20-SEP-2002
Mus musculus, similar to thymidylate kinase family LFS-inducible
member, clone IMAGE:3982036, mRNA.

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BC027329.1 GI:20072707

HTC

Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.

1 (bases 1 to 3124)

Direct Submission

Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LMNL)

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: angbcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 36 Row: e Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein

This clone has the following problem: frame shifted.

Location/Qualifiers

1. 3124

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

/db_xref="taxon:10090"

/clone="IMAGE:3982036"

/tissue_type="Mammary tumor metastasized to lung. Tumor

arose spontaneously from a senescent normal mammary

(clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"


```
QY 226 GTTGGGCCCCGGAGCGCAGCTATCTCGTGTGCGTGCCTGACCCCGGACGCCGGCTG 285
Db 343 GCTCGGCCCCCGAGGCGCAGCTACGCGTGTGTGCGCCCTGGCTCCAGGCGAAGGCTG 402
QY 286 CGGGGCCCCGGTCCGGGCGCGCGGTGACAGCGCCCTGCTGCACACGCTGCAGCGCGG 345
Db 403 CGGGGCCCCGGTGCAGGCGCGCGGTGCACATCGCCCTGCTGCACAGCTGCAGCGCGG 462
QY 346 CCGCTTCCAGCGGTGCAGCTGCTCAGGCTGCTCTGCTACTGCGCGGCGCGGCGAGCCCG 405
Db 463 TCTTTACAGAGATCCAACTGAGCAAGCTTCTGGGCTATGCTCGGGCGATCAGGCCCG 522
QY 406 CGGCGCACAGAAAGGCTCTGCTGTCGCGACACCCCTGGATGACCTGCACCCGCGCAAGC 465
Db 523 CGAAGCCACAGCTGGCTTCTGCTGCGGGACCCCTTGCACACCCCGGACACTCGGCGCGA 582
QY 466 GCTGCTCGAGCTGCTGGGCGCTGCAGGAGGACACCGCCGCACTTGGGCGAGTTGCA 525
Db 583 CTTGCTCCAGCTTCTGGCTCTCCAGAGGCGCGCGCCGCGAGTTGGCCGAGTTCCA 642
QY 526 GCGCGACCGCGCGGCGAGCTGTGGCAGCGCCCTCTCGGAGGTGCAAGCGCAGGCGGCT 585
Db 643 GCGCGACTCCAGGCTTCTGTGGCAGCGCTCTGTGGAGTTGCGAGGACAGCGCAGGT 702
QY 586 GCGGTGGGCTGCGACAGTCTGTCGCGCTCCCGGAGCCCGCTGCACCCGCTGCTGCC 645
Db 703 GCAGGTGGACTGCGCATGCGTCTGCGGACAGGAACTCTGTCACCCATGCTGCTGCC 762
QY 646 AGACTTGGCCAGTTCGCTGCTCTCCCGGACCGGAAAGCGCGCGGCGCTTTGGAGGA 705
Db 763 AGATCTGCTTAATCTGCGGTGTTCAGACCGGACCGGCGAGGCTCTTTGGAGA 822
QY 706 GTGTACTCTC-ATTATTCTTAAGCCCGGCGAGTGTGACCTGTGTCAGCATGCGCCAA 764
Db 823 GTGCACATCTTTATTCAGAAAGCCGGCGAGTACTTGACCTAGTTGACCATGCGCCAA 882
QY 765 AACAGATCCAGAAAGTTCAGGTTGTTGCCATCGAAGCACTGGATGCCACGGTA 824
Db 883 AGGAGTCCAGAAAGGAAGTTCCAGTCAITGCCATTGAAGGACTGGATGCCACTGGTA 942
QY 825 AAACCCAGGTGACCCAGTCACTGCGCAGATTCACCTTAAGGCTGCTCTTAAAGTCAAC 884
Db 943 AGACCACTGACGCACTGCTGCTGAGTCTCTCAAGGCTGCTCTACAGTGGCCAC 1002
QY 885 CCTCTTGCACTGCGAGTGGAGAGATCTTGTGATGATGAACCACTATCATTTAGAGAG 944
Db 1003 CCGCTGTATCAGCCAGTGGAGGAAGATCTTTGATGATGAACCTACTATCATTCGAAGAG 1062
QY 945 CTTTCTACTCTTTGGCAATTATATGTCGCTCCGAAATAGCTTAAAGATCTGCCAAAT 1004
Db 1063 CTTTCTACTCTTTGGCAATTATCTGTCGCTCTGAAATAGCTTAAAGATCTAACCAACT 1122
QY 1005 CTCCTGTGATTGAGACAGTACTTGGCAGACGCGCACCTATGCTCATAGCCACTGAGG 1064
Db 1123 TTCTGTATTGAGACAGTACTTGGCATAGCAGACGACCACTACGCTACTGAGG 1182
QY 1065 TGAGTGGGCTCTCCAGCACTGCGCCCGACCCATCACCTGTGTACAGTGGCCAGG 1124
Db 1183 TGAGTGGAGGCTACAGTACTACCTCCCGCCACCACTGTTGACAGTGGCCAGGG 1242
QY 1125 ACCTGCTCAACCTGACCTTATCTGCTGCTCACTGTGAGTCTCTGAGGAGGTTGAGA 1184
Db 1243 ACCTGCTGAAGCCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
QY 1185 GGTGTCAGGCGCGGCGATGAGAGACACGAGGAGAGCAGAACTTTGAGCCCAACAGT 1244
Db 1303 GACTGCGAGGCGCGGGTCCAGAGAAATCTAAAGAGAGGCTGAACTTTGAGCCCAATATG 1362
QY 1245 TGTTTCTCAAAAGTAGAATCTCTACAGCGGATGGAAATCTGCTGCTGCTGCTGCTGCTG 1304
Db 1363 TGTTTCTGAGAGGTTGAAATGAGTACCGACGCTATGGAGAACCCCAAGCTGCCATCTGG 1422
```

```
QY 1305 TTGATGCCAGCCCTCCAGAGAAAGGTCCTGCAGACAGTATTAAAGCTTAATCCAGATA 1364
Db 1423 TCGATGCCAGCCCTCCGAGAGAGACAGTCTCGAGAGGTTTATAGAGTGTATCCAGATT 1482
QY 1365 GTTTAGTGAACCGTAGTTACTCTGCGCAGGTGCCACGTCTAACTAGATAGATTGTT 1424
Db 1483 CTGTGCTGTTAATTAGTTGTTCTGCGCATGTTCTCTATTACACCGG--AGAAGCTGTG 1539
QY 1425 TGAACATCTACATCCACCATTTGTTATGCAAGTGTCCCAAAATTTCTGTTCTACAAGCAT 1484
Db 1540 TCGATGCCCGAAACCATCATCTGCGATCAG-CTTCTTAGATTAGGGGATGAGCAT 1598
QY 1485 GTTGTGTGCGCAGAAA 1500
Db 1599 ATTGTGTGAGAAAAA 1614

RESULT 4
AK004595 3022 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus adult male lung cDNA, RIKEN full-length enriched
DEFINITION library, clone:120004E04 product:thymidylate kinase family
IP5-inducible member, full insert sequence.
ACCESSION AK004595
VERSION AK004595.1 GI:12835870
KEYWORDS HTC: CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
3
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159
4
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwaga, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
5
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
REFERENCE
6
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
(bases 1 to 3022)
ADACHI, J., AIZAWA, K., AKAHIRA, S., AKIMURA, T., ARAI, A., AONO, H.,
ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FUKUNISHI, Y.,
FURUNO, M., HANGAKI, T., HARA, A., HAYATSU, N., HIRAMOTO, K.,
HIRAOKA, T., HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., IZAWA, M.,
```

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resesgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'. GAGAGAGAGCGCGCACTCGAGTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGGATCCAGGCTCAATTAATTAATTAACCCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES

source

Location/Qualifiers
 1..3022
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:1200004E04"
 /db_xref="taxon:10090"
 /clone="1200004E04"
 /sex="male"
 /tissue type="lung"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev stage="adult"
 129..1316
 /note="unnamed protein product; putative
 thymidylate kinase family LPS-inducible member
 (MGDI:G1:99830)"
 /codon_start=1
 /protein_id="BAB23396.1"
 /db_xref="GI:12835871"
 /translation="MQQTGMHAWKSCSAPQGAATRLVPLAPGCGGPRVQAAVHH
 RLQQLRRGFLQRCQLSLGLYGPQAGEAQLGELLRCDHPDTRDLQLGLSCQ
 EAARFLAEFQADSGLLMQRLWEIQGQVQVDCVLPAPGPHLHLLPLLSAV
 FQDRAARVLEKTSFIPEARVLDLVQCPKQVKGKQFVIAEGDATGKTLTQ
 SVSESLKAVLLSPPCISOWRKIFDEPTIIRAFYSIGNYLVAEIAKESTNPVI
 VDRYKHSATYALAEVSGLOLYLPAAHPVQWPCDILKPLVLLLTWNSEVRRL
 QRGGEKKEAELEANNVFRKVENTYORMENPSCHLVDSFRETVLQKVLLELIQS
 SGR"

CDS

polyA_signal
 3000..3005
 /note="putative"
 3022
 /note="putative"

ORIGIN

Query Match 33.9%; Score 872.4; DB 3; Length 3022;
 Best Local Similarity 77.1%; Pred. No. 4.4e-195;
 Matches 1119; Conservative 0; Mismatches 316; Indels 17; Gaps 4;
 QY 50 CTCCTCGGGGCACTGTGCGGGCGCGTGTGCTGGCGGGCGGGGTCTGCGCTGGGGCC 109
 DB 2 CTGTGCTCGGGGCTGTGCGGAGCGTGTGCGGGCGACACAGGCGCTGCATGGGGACC 61
 QY 110 ATGGCTCCCGCGCGCTTCGTCCTGGAGCTTCCGAGCTGCACCCCTGGCTCACTTCGCC 169

Db 62 ATGGCTCGGCCCGGGGGCTTCACTGTGAGCTGCCAGATTGCTCCCTGACTCACTTCGTC 121
 QY 170 CTAGGCGCGACGCCCCCGGCGACGACGCGCCCCCGACGCCCCCGCTGCGGGCGCTGTTG 229
 Db 122 CTGGG-----GGATCAACAGACACACAGGATGCAACGCTGCGAGCTGCTC 169
 QY 230 GGGCCCCCGGAGCGGAGCTACTCG-CTGTGCGTGCCTGCGTGCACCCCGAGCGCGGTGGG 288
 Db 170 GGGCCCCCAGGCGGAGCTACGCGCTGTGTGTGCGCTTGGCTCCAGGCGAAAGGCTGGCG 329
 QY 289 GGGCGGGTCCCGGCGCGCGCTGCACACAGCGCTGTGTCACACAGCTGCGCGCGGCGCC 348
 Db 230 GCGCGGGTGCAGGCGGCGCGGTGCACCATGCTGTGTCGACGAGCTGCGCGGGGTCC 289
 QY 349 CTTCCAGCGGTGCCAGCTGCTCAGGCTGCTGTGCTACTGTGCCGGGCGGCCAGCGCGGG 408
 Db 290 TTTACAGAGATGCCAACTGAGCAAGCTTCTGGGCTATGCTCCGGCGATCAGCGCGCGCA 349
 QY 409 CGCAGCAGNAGGCTTTCCTGCTGGCGACCCCTGTGATGACCTGCACCCGCGCAAGCGT 468
 Db 350 AGCCAGCATGCTTCTCTGCTGCGGACCCCTTGCACACACCCGCGACACTCGCGCGGACT 409
 QY 469 GCTCAGCTGTGCGGCGCTGCGAGGACACACGCGCGCACCTTGGGCGAGTTCGAGGC 528
 Db 410 GTTCCAGCTTCTGGGCTCTGCGAGGCGGCGCGCGCTGCTGGCCGAGTTCAGGC 469
 QY 529 CGACCGCGCGCGCAGCTGTGCGAGCGCTCTGGAGAGTGCAGAGCGGCGAGCGGTGCA 588
 Db 470 CGACTCCAGGCTTGTGCTGCGCAGCGCTGTGCGGAGTTCAGGGAGACAGCGCGGTGCA 529
 QY 589 GTTGGGCTGCGCACAGGTGTCGCGCTGCGGAGCGCGCTGCACCCCGTGGTGGCGCA 648
 Db 530 GGTGAGCTGCGCATGCTCTGCGCGCACAGAACTCATCTGTGACCCCTATGCTGCGCA 589
 QY 649 CTTGCCAGTTCGCTGGTCTTCCCGACCGGAAAGCGCGCGCGCTTTGGAGAGTG 708
 Db 590 TCTGCTTAACCTGCGGTGTTCCAGACCGGACCGGCGCAAGGCTGTCTTGAAGAGTG 649
 QY 709 TACCTCCTTTATTCCTGAAGCGCGGAGTGTGCTGACCTGTGCGTGCACCTGCGCCAAA 768
 Db 650 CACATCCTTTATTCAGAAAGCGCGGAGTCTTGACCTAGTTGACCTGCGCCAAAGGA 709
 QY 769 GATCCAGAAAGAAAGTTCAGGTTGTCCTCATCGAAGACTGGATGCCCGGTAAAC 828
 Db 710 GGTCCAGAAAGGAAAGTTCAGGTTGTCCTCATCGAAGACTGGATGCCCGGTAAAGAC 769
 QY 829 CACGCTGACCCAGTCACTGCGCAGATTCACCTTAAGGCTGTCTTAAAGTCAACACCTC 888
 Db 770 CACACTGACGAGTCACTGCGAGTCTCTCAAGGCTGTCTCTCTACAGTCCGACCCCC 829
 QY 889 TTGCATTTGCCAGTGGAGGAAGATCTTGATGATGAACCAATATCATTTAGAGAGCTTT 948
 Db 830 CTGTATCAGCCAGTGGAGGAAGATCTTTGATGATGAACCTACTATCATTCGAAGAGCTT 889
 QY 949 TTACTCTTTGGGCAATATATTGTGCTCCCAATAGTAAGATCTGCCAATCTCC 1008
 Db 890 TTATTCTTTGGGCAATATCTCTGCTTCTGAAATAGCTAAAGAAATCAACCACTTCC 949
 QY 1009 TGTGATTTAGACAGGTACTGCGACAGCAGCGCCACTATGCGATAGCCACTGAGGTGAG 1068
 Db 950 TGTATTGTAGACAGGTACTGCGATAGCAGCAGCCACTACGCCATAGCTACTGAGGTGAG 1009
 QY 1069 TGGGGGTCTCCAGCACCTGCGCCCCCGAGCCCATCACCTGTGTACCTAGTGCGCCAGAGCT 1128
 Db 1010 TGGAGGCTTACAGTACCTACCCCTGCCACCCCTGTGTACCTAGTGGCGGAGGACCT 1069
 QY 1129 GCTCAAACTGACCTTATCTGCTGCTCTGCTGCTGAGTCTGAGGAGAGGTTTCAGAGGCT 1188
 Db 1070 GCTGAAGCCGACTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1129
 QY 1189 GCAGGCGCGGCGCATGGAGAACACAGGGAAGAACAGCAAGCTTGAGGCCAACAGTGTGT 1248
 Db 1130 GCAGGCGCGGCGTCAGGAGAAACCTAAGAGAGGCTGAACTTGAGGCCAATATGTGT 1189


```

TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Faraday Avenue
REFERENCE  2 (bases 1 to 1287)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT    1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES   Location/Qualifiers
           1..1287
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DH005Y12"
             /tissue_type="T cells (Jurkat cell line)"
             /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      30.0%; Score 771.4; DB 3; Length 1287;
Best Local Similarity 98.5%; Pred. No. 2.9e-171;
Matches 852; Conservative 0; Mismatches 6; Indels 7; Gaps 7;
QY 1629 TAAACATTTCTTGAACATATGTTTCACTTAATCACTACCAATATCTGGAGACCTGT 1688
Db 1 TAAACATTTCTTGAACATATGTTTCACTTAATCACTACCAATATCTGGAGACCTGT 60
QY 1689 CTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGATCAGCAGG 1748
Db 61 CTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGATCAGCAGG 120
QY 1749 AGACCCGGAGGCTCTGCTTCTCCTACATGCTGATGATGAGATCGTGAATGCCAC 1808
Db 121 AGACCCGGAGGCTCTGCTTCTCCTACATGCTGATGATGAGATCGTGAATGCCAC 180
QY 1809 ATTGGCTTCTCCACATCTGTTTGCACTCGTATGATGGCTCGCTGATCTCCCTCAGT 1868
Db 181 ATTGGCTTCTCCACATCTGTTTGCACTCGTATGATGGCTCGCTGATCTCCCTCAGT 240
QY 1869 CCCAAATCTAGTAGCCTAGTCTTCTGTCAGAGGCTGTCTATGTCTCTGGTGCCTCAAG 1928
Db 241 CCCAAATCTAG-AGCCAAAGTCTTCTGTCAGAGGCTGTCTATGTCTCTGGTGCCTCAAG- 298
QY 1929 GGACACTCTGTCAGAGCCATTTTGGTAAGGAACACTTACAAAGAGGCATGTGATCTTG 1988
Db 299 GGACACTCTGTCAGAGCCATTTTGGTAAGGAACACTTACAAAGAGGCATGTGATCTTG 358
QY 1989 TGTCTGAGCTCAGAGCCCTTTTGATAGGCTTCTGATGTCATTCATTAAGACATTCAGC 2048
Db 359 TGTCTGAGCTCAGAGCCCTTTTGATAGGCTTCTGATGTCATTCATTAAGACATTCAGC 417
QY 2049 CAAGATGCTCCAAATGCAATATACCAACTTCTCTGAATATATTTTCTTATTATAT 2108
Db 418 CAAGATGCTCCAAATGCAATATACCAACTTCTCTGAATATATTTTCTTATTATAT 477
QY 2109 TTCTTTTCTTTTCTTAAAGATTTGGCTCTGATAGATGACATTTTCCATCTGAAT 2168
Db 478 TTCTTTTCTTTTCTTAAAGTA-TGGCTCTGAATAGATGACATTTTCCAT-TGAAT 535
QY 2169 GGATGCATATCATTTAGCAATCCAGTATTTTATTTATTTATTAATCATACATAATGTT 2228
Db 536 GGATGCATTTCAATTTAGCAATCCAGTATTTTATTTATTTATTAATCATACATAATGTT 595
QY 2229 TCCTCAGCATAGGAGTATGATTCATTAATTAAGTGGAGTCAAAACGCTAAATGCAAT 2288
Db 596 TCCTCAGCATAGGAGTATGATTCATTAATTAAGTGGAGTCAAAACGCTAAATGCAAT 655

```

us-10-681-223-1.rst

QY 2289 GTTCTTGTGATTTTCAATTACACAACTTAATTTTCTTCTTAAATAAGTTCAAGTGA 2348

Db 656 GTTCTTGTGATTTTCAATTACACAACTTAATTTTCTTCTTAAATAAGTAC-AGTGA 714

QY 2349 TCTTGGAGTGGGATTTCTTGGTAAATATCTTGGCACTTGAATGTCTCATGATTACATATG 2408

Db 715 TCTTGGAGTGGGATTTCTTGGTAAATATCTTGGCACTTGAATGTCTCATGATTACATATG 774

QY 2409 AAATCGCTTTGACATPATCTTTAGACAGAAAAAAGTAGCTGAGTGGGGGAAATATAGA 2468

Db 775 AAATCGCTTTGACATPATCTTTAGACAGAAAAAAGTAGCTGAGTGGGGGAAATATAGA 834

QY 2469 GCTTGTGACTTTAGGAGTAGCT 2493

Db 835 GC-TGTGTGACTTTAGGAGTAGGT 858

RESULT 7

CF272319 949 bp mRNA linear EST 13-AUG-2003

LOCUS AGENCOURT_15193947 NIH_MGC_192 Homo sapiens cDNA clone

DEFINITION IMAGE:30511964 5', mRNA sequence.

ACCESSION CF272319

VERSION CF272319.1 GI:33628231

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 949)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D. National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgabbs-remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Agencourt cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM571 row: h column: 21 High quality sequence stop: 582.

FEATURES Location/Qualifiers

1..949

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30511964"

/tissue_type="Brain glioblastoma"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_192"

/note="Vector: pExpress1; Site 1: SmaI; Site 2: NotI; The library was constructed by reverse transcription of 1 ug mRNA using the oligo dT primer GCGGCCGC(T)20 and an RNaseH + MMLV reverse transcriptase. Second strand synthesis was carried out by standard methods. The cDNA was size selected by agarose gel for > 1.2 kb, digested with Not I and directionally cloned into the vector Express-1 at the SmaI/NotI sites. DNA from the primary library was used for in vitro transcription from the T7 promoter to produce biotinylated RNA transcripts. These biotinylated transcripts, along with blocking oligos to the poly-A, multiple cloning site and primer regions, were hybridized with single stranded circles produced by phageid production from the primary library to a Cot value of 10-20. Streptavidin/phenol extraction was utilized to remove DNA:RNA hybrids leaving un-hybridized single

QY 1560 CCGTCATCAAGAAGTCCCTCTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTCTTC 1619
 DB 360 CCGTCATCAAGAAGTCCCTCTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTCTTC 419
 QY 1620 CTGGCATAGTAACAATTTCTTGGAAACATATGTTTCACCTTAATCATCTACCAATATCTGG 1679
 DB 420 CTGGCATAGTAACAATTTCTTGGAAACATATGTTTCACCTTAATCATCTACCAATATCTGG 479
 QY 1680 AAGACCTGTCTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGA 1739
 DB 480 AAGACCTGTCTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGA 539
 QY 1740 TCAGCAGGAGACCCGGGAGCCTCTGCTTCTCTACACTGGATGCTGATGAGATGCTGA 1799
 DB 540 TCAGCAGGAGACCCGGGAGCCTCTGCTTCTCTACACTGGATGCTGATGAGATGCTGA 599
 QY 1800 CATGCCACATTTGGCTTCTTCCACATCTGTTTGCACCTCGTCTCATGATGGCTCGCTGCATC 1859
 DB 600 CATGCCACATTTGGCTTCTTCCACATCTGTTTGCACCTCGTCTCATGATGGCTCGCTGCATC 659
 QY 1860 TCCCTCAGTCCCAATTTAGTAGCAGCAAGTGTCTCTGCGAGAGGCTGTCTATGTGCTGG 1919
 DB 660 TCCCTCAGTCCCAATTTAGTAGCAGCAAGTGTCTCTGCGAGAGGCTGTCTATGTGCTGG 719
 QY 1920 CTGCCCAAGGACATCTCTGCGAGAGCAATTTTGG 1954
 DB 720 TTGCCAAGGACA---TCCCTGCGAGAGCAATTTGGG 751

RESULT 9

BG497576

LOCUS

DEFINITION 601859469F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4073560 5', mRNA linear EST 27-MAR-2001

ACCESSION

BG497576

VERSION

BG497576.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 767)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/Drp

cDNA Library Prepared by: CLONTECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L10M322 row: i column: 17

High quality sequence stop: 760.

Location/Qualifiers

1..767

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4073560"

/tissue_type="adenocarcinoma"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC 60"

/note="Organ: prostate; Vector: pDNR-LTB (Clontech);

Site1: Sf11 (ggccgctcggcc); Site2: Sf11

(ggccattatggcc); Double-stranded cDNA was prepared from

cell line RNA. 5' and 3' adaptors were used in cloning as

follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and

3' adaptor sequence:

5'-ATTCTAGAGCGGAGCGCGGCATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.5

FEATURES

source

kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.7%; Score 685.2; DB 4; Length 767;
 Best Local Similarity 97.9%; Pred. No. 6.6e-151;
 Matches 735; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

QY 1525 ACTTCAGGCATCGTACCCTCTTCTGACATGATGGAGCCCGTCATCACAAGAGTCCCTCTCAT 1584
 DB 1 ACTTCAGGCATCGTACCCTCTTCTGACATGATGGAGCCCGTCATCACAAGAGTCCCTCTCAT 60
 QY 1585 CATGTTCCAGTCAGAGGCCAGCGATGCTTCTTCTGCGCATAGTAAACATTTCTTGGGA 1644
 DB 61 CATGTTCCAGTCAGAGGCCAGCGATGCTTCTTCTGCGCATAGTAAACATTTCTTGGGA 120
 QY 1645 ACATATGTTTCACTTAATCACTACCAATATCTGGAAGACCTGCTTACTCAGACAGCAC 1704
 DB 121 ACATATGTTTCACTTAATCACTACCAATATCTGGAAGACCTGCTTACTCAGACAGCAC 180
 QY 1705 CAGGTGTACAGAAGCAGACAGCAAGATCTTCAGATCAGCAGGAGACCCCGAGGCTCT 1764
 DB 181 CAGGTGTACAGAAGCAGACAGCAAGATCTTCAGATCAGCAGGAGACCCCGAGGCTCT 240
 QY 1765 GCTTCTCCTACACTGGCATGCTGATGAGATCTGACATGCCCCACATTTGGCTTCTTCCA 1824
 DB 241 GCTTCTCCTACACTGGCATGCTGATGAGATCTGACATGCCCCACATTTGGCTTCTTCCA 300
 QY 1825 TCTGTTGTCACCTGTCATGATGGGCTCGCTGATCTCCCTCAGTCCCAAAATCTAGTAGC 1884
 DB 301 TCTGTTGTCACCTGTCATGATGGGCTCGCTGATCTCCCTCAGTCCCAAAATCTAGTAGC 359
 QY 1885 CAAAGTGTTCCTGCAGAGGCTGTCTATGTCTGCTGCTGCCAAGGACACTCCTGCAGAG 1944
 DB 360 CAAAGTGTTCCTGCAGAGGCTGTCTATGTCTGCTGCTGCCAAGGACACTCCTGCAGAG 418
 QY 1945 CCATTTTGGTGAAGAACACTTACAAAGAGGCAATGATCTGTGCTGAGGCTCAGAG 2004
 DB 419 CCATTTTGGTGAAGAACACTTACAAAGAGGCAATGATCTGTGCTGAGGCTCAGAG 478
 QY 2005 CCCTTTTGTAGAGCTTCTGATGTCATTCATAAAGACATTCAGCCCAAGATGCTCCAACTG 2064
 DB 479 CCCTTTTGTAGAGCTTCTGA-GTCATATAAAGACATTCAGCCCAAGATGCTCCAACTG 537
 QY 2065 CAAATATACCAACCTTCTCTGAATATATTTTGGCTTATTTATATTTCTTTTCTTTTTC 2124
 DB 538 CAAATATACCAACCTTCTCTGAATATATTTTGGCTTATTTATTTCTTTTCTTTTTC 597
 QY 2125 TAAAGATTTGGCTCTGAATAGATGACATTTTCCATCTGAATCTGAATGATGCATATCATTA 2184
 DB 598 TAAAGTAT--GGCTCTGAATAGATGACATTTTCCATTTCCATTTGGAATGGAATCAATTTCA 655
 QY 2185 GCCAATCCAGTAATTTATTTATTTATTAATCTATACATAATATCTTCTCCTCAGCATAGGAC 2244
 DB 656 GCCAATCCAGTAATTTATTTATTTATTAATCTATACATAATATCTTCTCCTCAGCATAGGAC 715
 QY 2245 TATGATTCATTAATTAATAAGTGGAGTCAAAA 2275
 DB 716 TATGATTCATTAATTAATAAGTGGAGTCAAAA 746

RESULT 10

BG497758

LOCUS

DEFINITION

602543017F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4665484 5', mRNA sequence.

ACCESSION

BG497758

VERSION

BG497758.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

BG497758 759 bp mRNA linear EST 27-MAR-2001

602543017F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4665484 5', mRNA sequence.

ACCESSION BG497758

VERSION BG497758.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 759)
				NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
				Contact: Robert Strausberg, Ph.D. Email: cgapsb@mail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1CM1471 row: a column: 05 High quality sequence stop: 736.
FEATURES	source			
		1..759		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:4665484"		
		/tissue_type="adenocarcinoma"		
		/lab_host="DH10B (T1 phage-resistant)"		
		/clone_lib="NIH_MGC 60"		
		/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site1: SfiI (ggcgctcgggc); Site2: SfiI (ggccattatggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCAGGCGGCACATG-TT (30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."		
ORIGIN				
	Query Match	26.2%	Score 674.8; DB 4; Length 759;	
	Best Local Similarity	95.8%	Fred. No. 1.9e-148;	
	Matches 724; Conservative 0;	Mismatches 29; Indels 3; Gaps 3;		
Qy	1207	GAAGACACAGGGAAGAAGCAGAACTTGAGGCCAACAGTGTGTTTCGTCAAAGGTAGAAAT	1266	
Db	1	GAAGACACAGGGAAGAAGCAGAACTTGAGGCCAACAGTGTGTTTCGTCAAAGGTAGAAAT	60	
Qy	1267	GTCCTACAGCGGATGGAGATCTTGCTGCATGTGGTTGATGCCAGCCCTCCAGAGA	1326	
Db	61	GTCCTACAGCGGATGGAGATCTTGCTGCATGTGGTTGATGCCAGCCCTCCAGAGA	120	
Qy	1327	AAAGTCTCTGCAGACAGTATTAAAGCTTAATCCAGAAATAGTTTATGTGAAACCGTAGTTACT	1386	
Db	121	AAAGTCTCTGCAGACAGTATTAAAGCTTAATCCAGAAATAGTTTATGTGAAACCGTAGTTACT	180	
Qy	1387	CTGGCCAGGTGCCAGCTCTAATAGATTAGATGTTGTTGAAACATCTACATCCACCATT	1446	
Db	181	CTGGCCAGGTGCCAGCTCTAATAGATTAGATGTTGTTGAAACATCTACATCCACCATT	240	
Qy	1447	TGTTATGACAGTGTCCCAAAATTTCTGTGTTCTCAAGACATGTTGTGTGGCAGAAAATCTGGAG	1506	
Db	241	TGTTATGACAGTGTCCCAAAATTTCTGTGTTCTCAAGACATGTTGTGTGGCAGAAAATCTGGAG	300	
Qy	1507	ACCAGGCATCTTAATTTTACTTACGCCATCGTACCCCTCTTCTGACTGATGAGACCCGTCAT	1566	
Db	301	ACCAGGCATCTTAATTTTACTTACGCCATCGTACCCCTCTTCTGACTGATGAGACCCGTCAT	360	
Qy	1567	CACAAAAGTCCCTCTCATCATGTTTCCAGTGTAGAGCCAGCGATTGCTTTCTTCTGTGGCAT	1626	
Db	361	CACAAAAGTCCCTCTCATCATGTTTCCAGTGTAGAGCCAGCGATTGCTTTCTTCTGTGGCAT	420	
Qy	1627	AGTAAACATTTTCTTTGGAACATATGTTTTCACCTTAATCTACTACCAATATTTGGAGACCT	1686	

Db	421	AGTAAACATTTCTTGGACATATGTTTCACTTAATCACTACCAAAATATCTGGAAGACCT	480
Qy	1687	GTCTTACTCAGACAGCACACAGGTGTACAGAGCAGCAGACAAGATCTTCCAGATCAGCAG	1746
Db	481	GTCTTACTCAGACAGCACACAGGTGTACAGAGCAGCAGACAAGATCTTCCAGATCAGCAG	540
Qy	1747	GGAGACCCGGAGCCTCTGCTTCCTCCCTACACT--GGCATGCTGATGAGATCGTCACATGCC	1805
Db	541	GGAGACCCGGAGCCTCTGCTTCCTCCCTACACTGGGCATGCTGAAGAGATCGTGACATGCC	600
Qy	1806	CACATTTGGCTTCTTCCACATCTGGTTGCACCTCG--TCATGATGGGCTCGCTGCATCTCCCT	1864
Db	601	CACATTTGGCTTCTTCCACATCTGGTTGCACCTCGTTTATGATGGCTCGCTGCATCTCCCT	660
Qy	1865	CAGTCCCAAAATCTAGTAGCCAAAGTGTCTCTGCAGAGGCTGTCTATGTGTCTTGGCC--TGC	1923
Db	661	CAAGTCCCAAAATCTAGTAGCCAAAGTGTCTCTGCAGAGGCTGTCTATGTGTCTTGGCTTGC	720
Qy	1924	CCAAGGACACTCTCGCAGAGCATTTTTGGGTAAG	1959
Db	721	CCAAGGACACTCTCTGCAGAACCAATTTAGGGTAAG	756
RESULT 11			
BUL83182			
LOCUS			
DEFINITION			
BUL83182			
ACCESSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

ORIGIN

```
Query Match      25.4%; Score 653.2; DB 5; Length 769;
Best Local Similarity 98.6%; Pred. No. 2.5e-143;
Matches 680; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1267 GTCTACACGAGGATGAGAAATCTCGCTGCCATGTGGTTGATGCCAGCCCTCCAGAGA 1326
DB 3 GTCTACACGAGGATGAGAAATCTCGCTGCCATGTGGTTGATGCCAGCCCTCCAGAGA 62

QY 1327 AAAGTCTCTGACAGAGTATTAAAGCCCTAATCCAGAAATAGTTTGTAGTGAACCGTAGTTACT 1386
DB 63 AAAGTCTCTGACAGAGTATTAAAGCCCTAATCCAGAAATAGTTTGTAGTGAACCGTAGTTACT 122

QY 1387 CTGGCCAGGTGCCAGCTCTAACTAGATAGATGTTGTTGAAACATCTACATCCACATTT 1446
DB 123 CTGGCCAGGTGCCAGCTCTAACTAGATAGATGTTGTTGAAACATCTACATCCACATTT 182

QY 1447 TGTATGTCAGTGTTCCTCAAAATTTCTCTCTACAGCATGTGTGTGCGCAGAAAATGGAG 1506
DB 183 TGTATGTCAGTGTTCCTCAAAATTTCTCTCTACAGCATGTGTGTGCGCAGAAAATGGAG 242

QY 1507 ACAGGCACTCTTAATTTTACTTCAGCCATCGTATCCCTCTTCTGACTGTATGGACCCGTCAT 1566
DB 243 ACAGGCACTCTTAATTTTACTTCAGCCATCGTATCCCTCTTCTGACTGTATGGACCCGTCAT 302

QY 1567 CACAAAGGTCCCTCTCATCATGTTCAGTGAGAGGCCAGGATGCTTTCTTCTGGCAT 1626
DB 303 CACAAAGGTCCCTCTCATCATGTTCAGTGAGAGGCCAGGATGCTTTCTTCTGGCAT 362

QY 1627 AGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAACCT 1686
DB 363 AGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAACCT 422

QY 1687 GTCTTACTCAGACAGCACCAGGTGTACAGAGCAGCAGACAAGATCTTCCAGATCAGCAG 1746
DB 423 GTCTTACTCAGACAGCACCAGGTGTACAGAGCAGCAGACAAGATCTTCCAGATCAGCAG 482

QY 1747 GGAGACCCCGGAGCCTCTGCTCTCTACACATGCTGTGATGATGATGATGATGATGATGATG 1806
DB 483 GGAGACCCCGGAGCCTCTGCTCTCTACACATGCTGTGATGATGATGATGATGATGATGATG 542

QY 1807 ACATTGGCTTCTTCCACATCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1866
DB 543 ACATTGGCTTCTTCCACATCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602

QY 1867 GTCCCAAAATTCCTAGTACCAAGTGTTCCTGACAGAGCTCTCTATGTTGTTGTTGTTGTT 1925
DB 603 GTCCCAAAATTCCTAGTACCAAGTGTTCCTGACAGAGCTCTCTATGTTGTTGTTGTTGTT 661

QY 1926 AAGGACACTCTGCTGAGAGCCATTTTGGG 1955
DB 662 AAGGACACTCTGCTGAGAGCCATTTTGGG 691

RESULT 12
BG778349
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1655 row: m column: 21
High quality sequence stop: 701.
Location/Qualifiers
1. .705
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4806332"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 60"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggccgctcgccc); Site 2: SfiI
(ggccattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.5
kb (range 0.9-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      25.1%; Score 646.6; DB 4; Length 705;
Best Local Similarity 98.2%; Pred. No. 8.9e-142;
Matches 696; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

QY 1441 ACCATTGTGTATGCGAGTGTCCCAATTTCTGTTCTACAGCATGTTGTGCGCAGAAA 1500
DB 1 ACCATTGTGTATGCGAGTGTCCCAATTTCTGTTCTACAGCATGTTGTGCGCAGAAA 60

QY 1501 CTGGAGACAGGCACTTTAAATTTTACTTCAGGCATCTGACCTCTTCTGACTGATGACC 1560
DB 61 CTGGAGACAGGCACTTTAAATTTTACTTCAGGCATCTGACCTCTTCTGACTGATGACC 120

QY 1561 CGTCTATCAAAAGGTCCCTCTCATCATGTTCCAGTGAGAGGCCAGATGTTGTTCTTCC 1620
DB 121 CGTCTATCAAAAGGTCCCTCTCATCATGTTCCAGTGAGAGGCCAGATGTTGTTCTTCC 180

QY 1621 TGGCATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATCTGGA 1680
DB 181 TGGCATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATCTGGA 240

QY 1681 AGACCTGTCTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGACAAGATCTTCAGAT 1740
DB 241 AGACCTGTCTTACTCAGACAGCACCAGGTGTACAGAAGCAGCAGACAAGATCTTCAGAT 300

QY 1741 CAGCAGGAGAGACCCCGAGGCTCTGCTTCTCTA CACTGCGCATGCTGATGAGATCGTGAC 1800
DB 301 CAGCAGGAGAGACCCCGAGGCTCTGCTTCTCTA CACTGCGCATGCTGATGAGATCGTGAC 360

QY 1801 ATGCCCCACATTTGGCTTCTTCCACATCTGTTGTCACCTCGTCATGATGGGCTCGCTCATCT 1860
DB 361 ATGCCCCACATTTGGCTTCTTCCACATCTGTTGTCACCTCGTCATGATGGGCTCGCTCATCT 420

QY 1861 CCTCAGTCCCAAAATTTCTAGTAGCCAAGTGTTCCTGCGAGAGGCTCTCTATGTTGTTCTGGC 1920
DB 421 CCTCAGTCCCAAAATTTCTAGTAGCCAAGTGTTCCTGCGAGAGGCTCTCTATGTTGTTCTGGC 479

QY 1921 TGCCCAAGGAGACATCTCTGCGAGAGCCATTTTGGGTGAAGGAACACTTACAAAGAGGCAT 1980
DB 480 TGCCCAAGGAGACATCTCTGCGAGAGCCATTTTGGGTGAAGGAACACTTACAAAGAGGCAT 538

QY 1981 TGATCTTGTGTTGAGGCTCAGAGCCCTTTGATAGGCTTCTGATGCTCATTCATTAAGAC 2040
```



```

Plate: LNCMI655 row: d column: 22
High quality sequence stop: 631.
Location/Qualifiers
1. .844
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4806117"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC 60"
/site="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcgctggcc); Site_2: SfiI
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCCGAGGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.5
kb (range 0.9-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.2%; Score 621.8; DB 4; Length 844;
Best Local Similarity 98.9%; Pred. No. 6.9e-136;
Matches 626; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1018 AGACAGTACTGCGACAGACGCGCCACCTATGCGATAGCCACTGAGTGAGTGGGGTCT 1077
DB |||||
QY 1078 CAGACACTGCCCCCGACCCCATCACCTGTGTACAGTGCGCCAGAGACCTGCTCAAAACC 1137
DB |||||
QY 1138 TGACCTTATCCCTGCTGCTACTGTGAGTCTGTAGGAGAGGTTCCAGAGCTGAGGCGCG 1197
DB |||||
QY 1198 GGCATGAGAGAACACGAGGAGAAAGACAACTTGAGGCGCAACAGTGTGTTGCTCAAAA 1257
DB |||||
QY 1258 GGTAGAAATGCTCTACAGCGGATGAGAAATCCTGCTGCCATGTCGTTGATGCCAGGCC 1317
DB |||||
QY 1318 CTCCAGAGAAAGGTCTCCAGACAGTATTAAAGCCTAATCCAGAAAGTATTAGTGAACC 1377
DB |||||
QY 1378 GTAGTTACTCTGCGCAGGTGCCACGCTCTAACTAGATGTGTTGAAACATCTACA 1437
DB |||||
QY 1438 TCACACATTGTTATGAGTGTTCACAAATTTCTGTTCACAGCAATGTTGTGGCAGA 1497
DB |||||
QY 1498 AAACCTGGACCAAGGATCTTAATTTTACTTCAGCCATGTCACCTCTCTCTGACTGATGG 1557
DB |||||
QY 1558 ACCCGTCATCAAAAAGTCCCTCTCATCATGTGTTCCAGTGAGAGGCCAGCGATTGCTTCT 1617
DB |||||
QY 1618 TCCCTGCATAGTAAACATTTTCTTTGGAACATAT 1650
DB |||||
QY 1650 TCCCTGCATAGTAAACATTTTCTTTGGAACCAT 633

FEATURES
source
1. .720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4659044"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_59"
/site="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctggcc); Site_2: SfiI (ggcattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGGAGGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 23.9%; Score 613.8; DB 4; Length 720;
Best Local Similarity 96.1%; Pred. No. 5.2e-134;
Matches 693; Conservative 0; Mismatches 22; Indels 6; Gaps 6;

QY 1557 GACCCGTATCAAAAGGTCCTCTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTC 1616
DB 2 GACCCGTATCAAAAGGTCCTCTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTC 61
QY 1617 TTCTCTGGCATAGTAACATTTCTTGGACATATGTTTCACTTAATCACTACCAATATC 1676
DB 62 TTCTCTGGCATAGTAACATTTCTTGGACATATGTTTCACTTAATCACTACCAATATC 121
QY 1677 TGGAGACACCTGTTTACTCAGACAGCACCAGGTGTACAGAGCAGCAGCAAGATCTTCC 1736
DB 122 TGGAGACACCTGTTTACTCAGACAGCACCAGGTGTACAGAGCAGCAGCAAGATCTTCC 181
QY 1737 AGATCAGAGGAGACCCCGAGCCTCTGCTTCTCTACACTGGCATGCTGATGATCG 1796
DB 182 AGATCAGAGGAGACCCCGAGCCTCTGCTTCTCTACACTGGCATGCTGATGATCG 241
QY 1797 TGACATGCCACATTTGGCTTCTTCCACATCTCGTGTGCACTCGTATGAGGCTCGCTCG 1856

```


GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: October 22, 2004, 18:08:37 ; Search time 1219 Seconds
 (without alignments)
 10806.892 Million cell updates/sec

Title: US-10-681-223-1
 Perfect score: 2571
 Sequence: 1 cggggcggaggcggcgctcg.....aggaggggcttttgcaccc 2571

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2571	100.0	2571	14	US-10-277-032-1
2	2571	100.0	2571	16	US-10-681-223-1
3	2276	88.5	2856	15	US-10-305-810-6
4	2125.2	82.7	2645	13	US-10-044-090-720
5	1130.4	44.0	20966	14	US-10-277-032-3
6	1130.4	44.0	20966	16	US-10-681-223-3
7	1035.4	40.3	1579	16	US-10-264-237-104
8	303.8	11.8	507	9	US-09-867-701-2642
9	219.4	8.5	497	9	US-09-783-590-5564
10	149.8	5.8	603	16	US-10-240-425-827
11	126.2	4.9	390	9	US-09-783-590-11496
12	79.8	3.1	2118	15	US-10-156-761-7362
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 6, Appli
					Sequence 720, App
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 104, App
					Sequence 2642, Ap
					Sequence 5564, Ap
					Sequence 827, App
					Sequence 11496, A
					Sequence 7362, Ap

13	79.8	3.1	9025608	15	US-10-156-761-1	Sequence 1, Appl
14	76	3.0	1413	10	US-09-953-348-51	Sequence 51, Appl
15	76	3.0	1413	15	US-10-267-255-51	Sequence 51, Appl
16	76	3.0	53500	10	US-09-953-348-76	Sequence 76, Appl
17	76	3.0	53500	15	US-10-267-255-76	Sequence 76, Appl
18	75.8	2.9	667	17	US-10-767-701-4633	Sequence 4633, Ap
19	75	2.9	1092	17	US-10-767-701-9733	Sequence 9739, Ap
c 20	74	2.9	1860	15	US-10-156-761-1558	Sequence 1558, Ap
c 21	74	2.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
22	73	2.8	1437	15	US-10-156-761-5988	Sequence 5988, Ap
23	71.8	2.8	1041	17	US-10-437-963-8597	Sequence 8597, Ap
c 24	71.6	2.8	2561	9	US-09-976-740-48	Sequence 48, Appl
c 25	71.6	2.8	2561	13	US-10-023-529-48	Sequence 48, Appl
c 26	71.6	2.8	2561	13	US-10-023-523-48	Sequence 48, Appl
c 27	71.6	2.8	2561	16	US-10-616-187-48	Sequence 48, Appl
c 28	71.6	2.8	2561	16	US-10-671-242-48	Sequence 48, Appl
c 29	71	2.8	1388	17	US-10-767-701-14298	Sequence 14298, A
c 30	71	2.8	1734	15	US-10-156-761-6476	Sequence 6476, Ap
c 31	70	2.7	15738	15	US-10-329-079-46	Sequence 46, Appl
32	70	2.7	61944	15	US-10-329-079-34	Sequence 34, Appl
33	68.8	2.7	909	11	US-09-758-759-156	Sequence 34, Appl
34	68.8	2.7	109519	11	US-09-758-759-1	Sequence 156, App
c 35	68.6	2.7	3459	15	US-10-369-493-31552	Sequence 1, Appli
c 36	67.2	2.6	1618	15	US-10-104-047-1463	Sequence 31552, A
37	66.8	2.6	1946	17	US-10-437-963-49955	Sequence 1463, Ap
38	66.4	2.6	1843	17	US-10-437-963-22190	Sequence 49955, A
c 39	66.2	2.6	2520	10	US-09-873-367C-159	Sequence 22190, A
c 40	66	2.6	924	15	US-10-156-761-1189	Sequence 159, App
c 41	66	2.6	1521	15	US-10-156-761-3755	Sequence 1899, Ap
c 42	66	2.6	1725	17	US-10-324-967-35	Sequence 3755, Ap
c 43	65.6	2.6	763	17	US-10-437-963-49370	Sequence 35, Appl
c 44	65.6	2.6	3957	15	US-10-200-562-193	Sequence 49370, A
c 45	65.6	2.6	3957	15	US-10-237-551-193	Sequence 193, App

ALIGNMENTS

RESULT 1
 US-10-277-032-1
 ; Sequence 1, Application US/10277032
 ; Publication No. US20030087294A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ming-Hui WEI
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001305 DIV
 ; CURRENT APPLICATION NUMBER: US/10/277,032
 ; CURRENT FILING DATE: 2002-10-22
 ; PRIOR FILING DATE: 2001-10-31
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2571
 ; TYPE: DNA
 ; ORGANISM: HomoSapien
 ; US-10-277-032-1

Query Match	100.0%;	Score 2571;	DB 14;	Length 2571;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2571;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGGGGCGGAGGCGGCGTCTGAGCGCCCATGGCTTCGCCCGCGCGCTCTCTGGCGG 60		
Db	1	CGGGGCGGAGGCGGCGTCTGAGCGCCCATGGCTTCGCCCGCGCGCTCTCTGGCGG 60		
QY	61	GCCACTGTCTGGGCGCGCTCTCGGGCGGCGCGGGGTCTCGCTGGGGCCATGGCTCCGCC 120		
Db	61	GCCACTGTCTGGGCGCGCTCTCGGGCGGCGCGGGGTCTCGCTGGGGCCATGGCTCCGCC 120		
QY	121	GTGCCGCTCTGCTCTGGAGCTTCCGACCTGACCCCTGGCTCACTTGCCTAGGCCCGCA 180		

121 GTGCGCTTGGCTTGGAGCTTCCGACGTGCAACCTGGGTCACTTCCGCTTAGCGCCCGA 180
181 CGCCCCGGGCGACGACAGCGCCCCCGACCCCGCTTGGGGCGCTGTGGGGCCCCCGGA 240
181 CGCCCCGGGCGACGACAGCGCCCCCGACCCCGCTTGGGGCGCTGTGGGGCCCCCGGA 240
241 GCGAGCTACTCGCTGTGGGTGCGCTGTGACCCCGGACGCGGGTGGGGGCCCCGGGTCCG 300
241 GCGAGCTACTCGCTGTGGGTGCGCTGTGACCCCGGACGCGGGTGGGGGCCCCGGGTCCG 300
301 GGGGGCGCGCTGACACAGCGCTGCTGACACAGCTGCGCGCGGCGCCCTTCCAGCGGTG 360
301 GGGGGCGCGCTGACACAGCGCTGCTGACACAGCTGCGCGCGGCGCCCTTCCAGCGGTG 360
361 CAGCTGCTCAGGCTGCTGTCTACTGCGCGGCGCGCACAGCGCGCGCGCGCACAGCAAGG 420
361 CAGCTGCTCAGGCTGCTGTCTACTGCGCGGCGCGCACAGCGCGCGCGCGCACAGCAAGG 420
421 CTTCTGTGCGGACCCCTGTGATGACCTGTGACACCCCGCAAGCGCTGCTCGAGCTGCT 480
421 CTTCTGTGCGGACCCCTGTGATGACCTGTGACACCCCGCAAGCGCTGCTCGAGCTGCT 480
481 GGGCGCTGCGAGGACACCGCGCACTTTGGGGAGTTGAGGCGCGACCCGCGCGG 540
481 GGGCGCTGCGAGGACACCGCGCACTTTGGGGAGTTGAGGCGCGACCCGCGCGG 540
541 CCAGCTGTGCGAGCGCTTGGGAGGTGCAAGACGCGAGCGCGCTGCGAGTGGGCTGCGC 600
541 CCAGCTGTGCGAGCGCTTGGGAGGTGCAAGACGCGAGCGCGCTGCGAGTGGGCTGCGC 600
601 ACAGTGTGCGGCTGCGGAGCGCGCGCTGCGACCCCGTGGTGGCAGCTTGGCGATTC 660
601 ACAGTGTGCGGCTGCGGAGCGCGCGCTGCGACCCCGTGGTGGCAGCTTGGCGATTC 660
661 CGTGTCTTCCGACCGGGAAGCGCGCGCGGCGTTTGGAGAGTGTACCTCTCTTTAT 720
661 CGTGTCTTCCGACCGGGAAGCGCGCGCGGCGTTTGGAGAGTGTACCTCTCTTTAT 720
721 TCCTGAAGCCGGGCGAGTGTGACCTGGTGCACGAGTGGCCAAAACAGATCCGAAAGG 780
721 TCCTGAAGCCGGGCGAGTGTGACCTGGTGCACGAGTGGCCAAAACAGATCCGAAAGG 780
781 AAAGTTCAGGTTGTTGCCATCGAAGACTGATGACCGGTAACCGGTAACCGGTAACCG 840
781 AAAGTTCAGGTTGTTGCCATCGAAGACTGATGACCGGTAACCGGTAACCGGTAACCG 840
841 GTCAGTGGCAGATTCACTTAAGGCTGCTCTTAAAGTCAACCCCTCTTGCATTGGCCA 900
841 GTCAGTGGCAGATTCACTTAAGGCTGCTCTTAAAGTCAACCCCTCTTGCATTGGCCA 900
901 GTGAGGAAGATCTTTGATGATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGG 960
901 GTGAGGAAGATCTTTGATGATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGG 960
961 CAATATATTTGCGCTCGGAAATAGTAAAGATCTGCCAAATCTCTGTGATTTAGA 1020
961 CAATATATTTGCGCTCGGAAATAGTAAAGATCTGCCAAATCTCTGTGATTTAGA 1020
1021 CAGTACTGGCAGCAGCGGCACTATGCCATAGCCACTGAGGTGAGTGGGGGTCTCCA 1080
1021 CAGTACTGGCAGCAGCGGCACTATGCCATAGCCACTGAGGTGAGTGGGGGTCTCCA 1080
1081 GCACCTGCCCCAGCCCATACCTGTGTACAGTGGCGCAGAGCTGCTCAACCTGA 1140
1081 GCACCTGCCCCAGCCCATACCTGTGTACAGTGGCGCAGAGCTGCTCAACCTGA 1140
1141 CTTATCTGTCTGCTCACTGTGAGTCCCTGAGGAGGTTGAGAGGCTGCGAGGCGCGGG 1200
1141 CTTATCTGTCTGCTCACTGTGAGTCCCTGAGGAGGTTGAGAGGCTGCGAGGCGCGGG 1200
1201 CATGAGAGACCGAGGGAAGACGAACTTGAGGCCAACAGTGTGTTTCTGTCAAAAGGT 1260

1201 CATGAGAGACCGAGGGAAGACGAACTTGAGGCCAACAGTGTGTTTCTGTCAAAAGGT 1260
1261 AGAATGTCTACACGCGGATGAGAACTCTCGCTGCCATGTGGTTGATGCCAGCCCCCTC 1320
1261 AGAATGTCTACACGCGGATGAGAACTCTCGCTGCCATGTGGTTGATGCCAGCCCCCTC 1320
1321 CAGAGAAAAGGTCCTGCGACAGTAGTAATAAGCCCTAAATCCAGAAATAGTTTAGTGAACCGTA 1380
1321 CAGAGAAAAGGTCCTGCGACAGTAGTAATAAGCCCTAAATCCAGAAATAGTTTAGTGAACCGTA 1380
1381 GTTACTCTGCGAGGTCACAGTCTAACTAGATAGATGTTGTTGAAAACATCTATATCC 1440
1381 GTTACTCTGCGAGGTCACAGTCTAACTAGATAGATGTTGTTGAAAACATCTATATCC 1440
1441 ACCATTGTATGACAGTGTCCCAAATTTCTGTTCTACAGCATGTTGTTGCGCAGAAA 1500
1441 ACCATTGTATGACAGTGTCCCAAATTTCTGTTCTACAGCATGTTGTTGCGCAGAAA 1500
1501 CTGAGACACGAGCATCTTAATTTTACTTCAGCCATCGTACCCCTCTCTGACTGATGGACC 1560
1501 CTGAGACACGAGCATCTTAATTTTACTTCAGCCATCGTACCCCTCTCTGACTGATGGACC 1560
1561 CGTCTACAAAGGTCCTCTCATCATGTTCCAGTGAGGCGCAGCATGCTTCTTCC 1620
1561 CGTCTACAAAGGTCCTCTCATCATGTTCCAGTGAGGCGCAGCATGCTTCTTCTCC 1620
1621 TGGCATAGTAAACATTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGA 1680
1621 TGGCATAGTAAACATTTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGA 1680
1681 AGACCTGTCTTACTCAGACAGCACCGAGGTGTAAGAGAGCAGCAGAAATCTTCCAGAT 1740
1681 AGACCTGTCTTACTCAGACAGCACCGAGGTGTAAGAGAGCAGCAGAAATCTTCCAGAT 1740
1741 CAGAGGGAGACCCCGGAGCCTCTGCTTCTCTACCTGCGCATGCTCATGAGATCTGAC 1800
1741 CAGAGGGAGACCCCGGAGCCTCTGCTTCTCTACCTGCGCATGCTCATGAGATCTGAC 1800
1801 ATGCCACATTTGGCTTCTTCCACATCTGTTGCACTCGTCACTGATGAGTCTGATCT 1860
1801 ATGCCACATTTGGCTTCTTCCACATCTGTTGCACTCGTCACTGATGAGTCTGATCT 1860
1861 CCCTCAGTCCCAAATCTAGTAGCCAAAGTGTCTCTGAGAGGCTGTCTATGTTGCTGGC 1920
1861 CCCTCAGTCCCAAATCTAGTAGCCAAAGTGTCTCTGAGAGGCTGTCTATGTTGCTGGC 1920
1921 TGCCCAAGGACACCTCTGCGAGGCAATTTTGGGTAGGAACTTACAAAGAGGCAT 1980
1921 TGCCCAAGGACACCTCTGCGAGGCAATTTTGGGTAGGAACTTACAAAGAGGCAT 1980
1981 TGATCTTGTGCTGAGGCTCAGAGCCCTTTGATAGGCTTCTGATGCTATCATAAGAC 2040
1981 TGATCTTGTGCTGAGGCTCAGAGCCCTTTGATAGGCTTCTGATGCTATCATAAGAC 2040
2041 ATTCAAGCAAGATGCTCCAACTGCAAAATATACCACTTCTCTGAAATATATTTGGTT 2100
2041 ATTCAAGCAAGATGCTCCAACTGCAAAATATACCACTTCTCTGAAATATATTTGGTT 2100
2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGCGCTGTAAGTGAATGCAATTTTCCA 2160
2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGCGCTGTAAGTGAATGCAATTTTCCA 2160
2161 TCTGAACCTGGAATGATATCAATTTAGCCAAATCCAGTAAATTTATTTATATATATATACAT 2220
2161 TCTGAACCTGGAATGATATCAATTTAGCCAAATCCAGTAAATTTATTTATATATATATACAT 2220
2221 AATATGTTTCTCAGCATAGGATGATCAATTAATAAGTGAAGTCAAAACGCTA 2280
2221 AATATGTTTCTCAGCATAGGATGATCAATTAATAAGTGAAGTCAAAACGCTA 2280
2281 AATGCAATGTTTGTGATTTTCAATACCAAACTTAAATTTGTTGTTTAAATAGTT 2340
2281 AATGCAATGTTTGTGATTTTCAATACCAAACTTAAATTTGTTGTTTAAATAGTT 2340

```

QY 2341 CAAGTGGATCTTGGAGTGGGATTTCTTGGTAAATTAATCTTGACATTTGAATGCTCATGAT 2400
Db 2341 CAAGTGGATCTTGGAGTGGGATTTCTTGGTAAATTAATCTTGACATTTGAATGCTCATGAT 2400
QY 2401 TACATATGAATCGCTTTGACATATCTTTAGACAGAAAAAAGTAGCTGAGTGAGGGGAA 2460
Db 2401 TACATATGAATCGCTTTGACATATCTTTAGACAGAAAAAAGTAGCTGAGTGAGGGGAA 2460
QY 2461 ATTATAGACTGTGTGATCTTTAGGAGTAGCTGTCTTTATACACATACTCAAGCCCTG 2520
Db 2461 ATTATAGACTGTGTGATCTTTAGGAGTAGCTGTCTTTATACACATACTCAAGCCCTG 2520
QY 2521 AAGCCTTGATGCTCGTGGAGCTCGACATAAGGAGGGGCTTTTGACCC 2571
Db 2521 AAGCCTTGATGCTCGTGGAGCTCGACATAAGGAGGGGCTTTTGACCC 2571

RESULT 2
US-10-681-223-1
; Sequence 1, Application US/10681223
; Publication No. US20040081999A1
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305 DIV-II
; CURRENT APPLICATION NUMBER: US/10/681,223
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 10/277,032
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-681-223-1

Query Match 100.0%; Score 2571; DB 16; Length 2571;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGGCGGAGGCGGCGTCTGAGAGCGCCCATGGCTTTGCGCGCGGCGCTCTGCGCGG 60
Db 1 CGGGGCGGAGGCGGCGTCTGAGAGCGCCCATGGCTTTGCGCGCGGCGCTCTGCGCGG 60
QY 61 GCACCTGTGCGGCGCGCTCTGCGGCGGCGGCGGCTCTGCGCTGGGCGCATGGCTCGGC 120
Db 61 GCACCTGTGCGGCGCGCTCTGCGGCGGCGGCGGCTCTGCGCTGGGCGCATGGCTCGGC 120
QY 121 GTGCGGCTCTGCTGAGACTTCCGACTGCACTTGGCTGCTTCTGCGCTTAGGCGCGGA 180
Db 121 GTGCGGCTCTGCTGAGACTTCCGACTGCACTTGGCTGCTTCTGCGCTTAGGCGCGGA 180
QY 181 CGCCCCGCGCAGCAGACGCCCGGACCCCGGCTGGGCGGCTGTTGGGCGCCCCCGGA 240
Db 181 CGCCCCGCGCAGCAGACGCCCGGACCCCGGCTGGGCGGCTGTTGGGCGCCCCCGGA 240
QY 241 GCGCAGCTACTGCTGTGCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCTCGG 300
Db 241 GCGCAGCTACTGCTGTGCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCTCGG 300
QY 301 GCGGCGGCGGCTGCAACAGCGCTGCTGCAACAGCTGCGCGGCGGCGGCTTCCAGCGGTG 360
Db 301 GCGGCGGCGGCTGCAACAGCGCTGCTGCAACAGCTGCGCGGCGGCGGCTTCCAGCGGTG 360
QY 361 CCAGCTGTCTCAGGCTGCTCTGCTACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 361 CCAGCTGTCTCAGGCTGCTCTGCTACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

```

```

QY 421 CTTCTGCTGCGGACGACCCCTCGATGACCTGACACCCCGGCAAGCGCTGCTCGAGCTGCT 480
Db 421 CTTCTGCTGCGGACGACCCCTCGATGACCTGACACCCCGGCAAGCGCTGCTCGAGCTGCT 480
QY 481 GGGCGCTGCCAGGAGGACACCGCGGCTGTTGGCGAGTTTCGAGGCGGCGGCGGCGG 540
Db 481 GGGCGCTGCCAGGAGGACACCGCGGCTGTTGGCGAGTTTCGAGGCGGCGGCGGCGG 540
QY 541 CCAGCTGTGCGGACGCGCTCTGCGGAGGTGCAAGACGCGGCGGCTGCGAGTGGGCTGCGC 600
Db 541 CCAGCTGTGCGGACGCGCTCTGCGGAGGTGCAAGACGCGGCGGCTGCGAGTGGGCTGCGC 600
QY 601 ACAGTGTGCGGCTGCGGAGCGCGCTGCGGAGCGCGCTGCGGAGCTGCGGAGCTGCG 660
Db 601 ACAGTGTGCGGCTGCGGAGCGCGCTGCGGAGCGCGCTGCGGAGCTGCGGAGCTGCG 660
QY 661 CGTGTGCTTCCCGGACCGGGAAGCGCGGCTGTTTGGAGGAGTGTACCTCCTTTAT 720
Db 661 CGTGTGCTTCCCGGACCGGGAAGCGCGGCTGTTTGGAGGAGTGTACCTCCTTTAT 720
QY 721 TCCTGAAGCCCGGCGGAGTGTGACCTGCTGACGAGTGCACCAAGTGCACCAACAGATCCAGAAAG 780
Db 721 TCCTGAAGCCCGGCGGAGTGTGACCTGCTGACGAGTGCACCAAGTGCACCAACAGATCCAGAAAG 780
QY 781 AAAGTTCAGGTTGTTGCCATCGAAGGACTGATGCGACGGGTAAACACCGGTGACCCA 840
Db 781 AAAGTTCAGGTTGTTGCCATCGAAGGACTGATGCGACGGGTAAACACCGGTGACCCA 840
QY 841 GTCAGTGGCAGATTCACATTAAGGCTGCTCTTAAAGTCAACACCTCTTGCATTTGGCCA 900
Db 841 GTCAGTGGCAGATTCACATTAAGGCTGCTCTTAAAGTCAACACCTCTTGCATTTGGCCA 900
QY 901 GTGAGGAGAGTCTTTGATGATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGG 960
Db 901 GTGAGGAGAGTCTTTGATGATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGG 960
QY 961 CAATTATATTGCGGCTCGGAAATAGCTAAAGAAATCTGCCAAATCTCTCTGTGATTGATA 1020
Db 961 CAATTATATTGCGGCTCGGAAATAGCTAAAGAAATCTGCCAAATCTCTCTGTGATTGATA 1020
QY 1021 CAGGTACTGGCACAGCAGCGCCACCTATGCCATAGCCTAGGAGTGGGCTCTCCA 1080
Db 1021 CAGGTACTGGCACAGCAGCGCCACCTATGCCATAGCCTAGGAGTGGGCTCTCCA 1080
QY 1081 GCACCTGCGCCCGAGCCCATCACCTGTGTACAGTGGCCAGAGGACCTGCTCAACCTGA 1140
Db 1081 GCACCTGCGCCCGAGCCCATCACCTGTGTACAGTGGCCAGAGGACCTGCTCAACCTGA 1140
QY 1141 CTTATCTCTGCTGCTCACTGTGAGTCTCTGAGAGAGTTCGAGGCTGCGAGGCGGCGG 1200
Db 1141 CTTATCTCTGCTGCTCACTGTGAGTCTCTGAGAGAGTTCGAGGCTGCGAGGCGGCGG 1200
QY 1201 CATGAGAGAGACCGGGAAGAGAGAACTTCAGAGCCCAACAGTGTGTTTCGTCAAAAGT 1260
Db 1201 CATGAGAGAGACCGGGAAGAGAGAACTTCAGAGCCCAACAGTGTGTTTCGTCAAAAGT 1260
QY 1261 AGAAATGTCTACACGCGGATGGAGAACTCTGCGCTGCCATGTGTTGATGCGAGCCCTTC 1320
Db 1261 AGAAATGTCTACACGCGGATGGAGAACTCTGCGCTGCCATGTGTTGATGCGAGCCCTTC 1320
QY 1321 CAGAGAAAGTCTCTGCGAGACAGTAAAGCCTAATCCAGAAATAGTTTTAGTGAACCGTA 1380
Db 1321 CAGAGAAAGTCTCTGCGAGACAGTAAAGCCTAATCCAGAAATAGTTTTAGTGAACCGTA 1380
QY 1381 GTTACTCTGGCGAGTGCACGCTAACTAGATTAGATGTTGTTTGAACATCTACATCC 1440
Db 1381 GTTACTCTGGCGAGTGCACGCTAACTAGATTAGATGTTGTTTGAACATCTACATCC 1440
QY 1441 ACCATTGTTATGACAGTGTCCCAAATTTCTGTTCTACAAGCATGTTGTGTGGCAGAAA 1500
Db 1441 ACCATTGTTATGACAGTGTCCCAAATTTCTGTTCTACAAGCATGTTGTGTGGCAGAAA 1500

```

1501 CTGGAGACGAGCATCTTAATTTTACTTCAGGCATCGTACCTCTTCTGACTGATGGACC 1560
1501 CTGGAGACGAGCATCTTAATTTTACTTCAGGCATCGTACCTCTTCTGACTGATGGACC 1560
1561 CGTCATCAAAAGGTCCTCTCATCATGTTCCAGTGAGAGGCGAGGATGCTTCTTCC 1620
1561 CGTCATCAAAAGGTCCTCTCATCATGTTCCAGTGAGAGGCGAGGATGCTTCTTCC 1620
1621 TGGCATAGTAAACATTTCTTGGACATATGTTTCACTTAATCACTPACCAATATCTGA 1680
1621 TGGCATAGTAAACATTTCTTGGACATATGTTTCACTTAATCACTPACCAATATCTGA 1680
1681 AGACCTGCTTACTCAGACAGACACGAGTGTACAGACAGACAGACAGATCTTCCAGAT 1740
1681 AGACCTGCTTACTCAGACAGACACGAGTGTACAGACAGACAGACAGATCTTCCAGAT 1740
1741 CAGCAGGAGACCCCGGAGCCTCTGCTTCTCTACACTGCGCATGCTGATGATCGTGAC 1800
1741 CAGCAGGAGACCCCGGAGCCTCTGCTTCTCTACACTGCGCATGCTGATGATCGTGAC 1800
1801 ATGCCACATGCGCTTCTTCCACATCTGGTTGACCTGCTGATGATGCTGCTGATCT 1860
1801 ATGCCACATGCGCTTCTTCCACATCTGGTTGACCTGCTGATGATGCTGCTGATCT 1860
1861 CCTCAGTCCCAATTTCTAGTACGCAAGTCTTCTGAGAGGCTGTCTATGCTCTGCGC 1920
1861 CCTCAGTCCCAATTTCTAGTACGCAAGTCTTCTGAGAGGCTGTCTATGCTCTGCGC 1920
1921 TGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAGGAACACTTACAAAGAGGCAT 1980
1921 TGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAGGAACACTTACAAAGAGGCAT 1980
1981 TGATCTGTGCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTAAGAC 2040
1981 TGATCTGTGCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTAAGAC 2040
2041 ATTCAAGCAGAGTGTCCCAACTGCAATATACCACTCTGCTGATTAATATTGCTT 2100
2041 ATTCAAGCAGAGTGTCCCAACTGCAATATACCACTCTGCTGATTAATATTGCTT 2100
2101 ATTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGACATTTTCCA 2160
2101 ATTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGACATTTTCCA 2160
2161 TCTGAATGATGATGATATCATTTAGCAATCCAGTAATTTATATTAATCTATACAT 2220
2161 TCTGAATGATGATGATATCATTTAGCAATCCAGTAATTTATATTAATCTATACAT 2220
2221 AATATGTTCTCAGCATAGAGCTATGATTCATTAATTAAGTGAAGTCAAAACGCTA 2280
2221 AATATGTTCTCAGCATAGAGCTATGATTCATTAATTAAGTGAAGTCAAAACGCTA 2280
2281 AATGCAATGTTTGTGTATTTTCAITACAAACTTAATTTCTGTTGTTAAATAAGTT 2340
2281 AATGCAATGTTTGTGTATTTTCAITACAAACTTAATTTCTGTTGTTAAATAAGTT 2340
2341 CAAGTGATCTTGAGTGGGATTTCTGTTGTTAAATTTCTGCACTGAAATGCTCATGAT 2400
2341 CAAGTGATCTTGAGTGGGATTTCTGTTGTTAAATTTCTGCACTGAAATGCTCATGAT 2400
2401 TACATATGAATCGCTTTGACATATCTTTAGACAGAAAAGTGTAGCTGAGTGGGGGAA 2460
2401 TACATATGAATCGCTTTGACATATCTTTAGACAGAAAAGTGTAGCTGAGTGGGGGAA 2460
2461 ATTATAGAGCTGTGTGACTTTAGGAGTAGCTGTCTCTTATACATATCTCAAGCCCTG 2520
2461 ATTATAGAGCTGTGTGACTTTAGGAGTAGCTGTCTCTTATACATATCTCAAGCCCTG 2520
2521 AAGCCTTGCATGCTCTGAGCGTGCACATAAGAGGGGGCTTTTGACCC 2571
2521 AAGCCTTGCATGCTCTGAGCGTGCACATAAGAGGGGGCTTTTGACCC 2571

RESULT 3
US-10-305-810-6
; Sequence 6, Application US/10305810
; Publication No. US20030176385A1
; GENERAL INFORMATION:
; APPLICANT: Ju, Jingfang
; APPLICANT: Huang, Chunli
; APPLICANT: Zhong, Haihong
; APPLICANT: Simons, Jan Fredrik
; APPLICANT: Rallou, Bruce E.
; APPLICANT: Chant, John S.
; APPLICANT: Peyman, John A.
; APPLICANT: Smithson, Glenna
; APPLICANT: Millet, Isabelle
; TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN EXPRESSION
; FILE REFERENCE: 21402-501
; CURRENT APPLICATION NUMBER: US/10/305,810
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/334,148
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/336,572
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/625,634
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/192,838
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/194,256
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/970,813
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/182,637
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/240,316
; PRIOR FILING DATE: 2000-10-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 6
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Thymidine Kinase
US-10-305-810-6
Query Match 88.5%; Score 2276; DB 15; Length 2856;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2386; Conservative 0; Mismatches 10; Indels 16; Gaps 9;
Qy 83 GGGCGGCGGGGCTCGCTGGGGCCATGGCTCCGCGCTCCGCTTTCGCTCGAGCTT 142
Db 1 GGGCGGCGGGGCTCGCTGGGGCCATGGCTCCGCGCGCCGCTTCGCTCGAGCTT 60
Qy 143 CCGGACTGCACTCGCTGCTTCCGCTTAGGGCCGCGCCGCGCGCGCGAGCGCC 202
Db 61 CCGGACTGCACTCGCTGCTTCCGCTTAGGGCCGCGCCGCGCGCGAGCGCC 120
Qy 203 CCGGACCCCGCTGGGGCGGCTTGGGGCCCGCGAGCGAGCTACTCGCTGGGTG 262
Db 121 CCGGACCCCGCTGGGGCGGCTTGGGGCCCGCGAGCGAGCTACTCGCTGGGTG 180
Qy 263 CCGGACCCCGCGAGCGCGCTGGGGCCCGCGGCGCGCGCGCGCGCTGACAGCGC 322
Db 181 CCGGAGCCCGCGAGCGCGCTGGGGCCCGCGGCGCGCGCGCGCGCTGACAGCGC 240
Qy 323 CTGCTGACCAAGCTGCGCGCGGCCCTTCCAGCGGTGCGAGCTGCTCAGCTGCTGCG 382
Db 241 CTGCTGACCAAGCTGCGCGCGGCCCTTCCAGCGGTGCGAGCTGCTCAGCTGCTGCG 300
Qy 383 TACTGCGGCTG 442

Db 301 TACTGCCGGGCGCCAGGCGCGGCGCACACAGCAAGGCTTCTGCTGCGGACCCCTG 360
QY 443 GATGACCTGACACCGCGGAGCGCTGCTGAGACTGCTGGCGCCTGCGCAGAGGACCA 502
Db 361 GATGACCTGACACCGCGGAGCGCTGCTGAGCTGCTGGCGCCTGTCAGAGGACCA 420
QY 503 CGCCCGCACTTGGCGAGTTCAGAGCGCGACCGCGCGGCGAGCTGCGCAGCGCTG 562
Db 421 CGCCCGCACTTGGCGAGTTCAGAGCGCGACCGCGCGGCGAGCTGCGCAGCGCTG 480
QY 563 GAGGTCAAGACGCGGCGGCTGCGAGGTGGGTGGGCGACAGGTGCTGCGCGCGGAG 622
Db 481 GAGGTCAAGACGCGGCGGCTGCGAGGTGGGTGGGCGACAGGTGCTGCGCGCGGAG 540
QY 623 CCCCCTGACCCCGTGTGCGAGACTTGCAGTTCGCCAGTTCCTGTCCTTCCGCGAG 682
Db 541 CCCCCTGACCCCGTGTGCGAGACTTGCAGTTCGCCAGTTCCTGTCCTTCCGCGAG 600
QY 683 GCGCCCGGCGCTTTTGGAGGAGTGTACCTCTCTTTATTTCTGAAGCCGCGAGTCT 742
Db 601 GCGCCCGGCGCTTTTGGAGGAGTGTACCTCTCTTTATTTCTGAAGCCGCGAGTCT 660
QY 743 GACCTGCTGACACGAGTCCCAACAGATCCAGAAAGAAAGTTCAGGTTGTGCGATC 802
Db 661 GACCTGCTGACACGAGTCCCAACAGATCCAGAAAGAAAGTTCAGGTTGTGCGATC 720
QY 803 GAAGGACTGATGCGAGCG---GGTAAACACCGTGACCGGAGTTCAGTTCAGTTC 859
Db 721 GAAGGACTGATGCGAGCGGTGTAACACCGGTGACCGGAGTTCAGTTCAGTTC 780
QY 860 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTGGCGAGTGGAGGAAGATCTTTGAT 919
Db 781 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTGGCGAGTGGAGGAAGATCTTTGAT 840
QY 920 GATGAACCAACTATCAATTAGAGAGCTTTTACTCTTTGGGCAATTTATTTGTGGCTCC 979
Db 841 GATGAACCAACTATCAATTAGAGAGCTTTTACTCTTTGGGCAATTTATTTGTGGCTCC 900
QY 980 GAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTGTAGACAGGTACTGGCAGCAG 1039
Db 901 GAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTGTAGAC-----GGCAGCAGCAG 954
QY 1040 GCCACCTATGCCATAGCCACTGAGTGGTGGGGTCTCCAGCACCTGCCCGAGCCCAT 1099
Db 955 GCCACCTATGCCATAGCCACTGAGTGGTGGGGTCTCCAGCACCTGCCCGAGCCCAT 1014
QY 1100 CACCTGTGTACAGTGGCGCAGAGGACCTGCTCAAACTGACCTTATCTGCTGTCACT 1159
Db 1015 CACCTGTGTACAGTGGCGCAGAGGACCTGCTCAAACTGACCTTATCTGCTGTCACT 1074
QY 1160 GTGAGTCTGAGGAGGTTGCAGAGCTGCGAGGCGCGGGCGGCGATGGAGAAGCAGGGAA 1219
Db 1075 GTGAGTCTGAGGAGGTTGCAGAGCTGCGAGGCGCGGGCGGCGATGGAGAAGCAGGGAA 1134
QY 1220 GAAGCAGAACTTGAGGCCAACAGTGTGTTTGTCAAAAGTGAATGCTTACAGCGG 1279
Db 1135 GAAGCAGAACTTGAGGCCAACAGTGTGTTTGTCAAAAGTGAATGCTTACAGCGG 1194
QY 1280 ATGGAGAATCTGGCTGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGTCTGTCAG 1339
Db 1195 ATGGAGAATCTGGCTGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGTCTGTCAG 1254
QY 1340 ACAGTATTAAAGCTAATCCAGATAGTTTTAGTGAACCGTAGTACTCTGGCCAGGTGCC 1399
Db 1255 ACAGTATTAAAGCTAATCCAGATAGTTTTAGTGAACCGTAGTACTCTGGCCAGGTGCC 1314
QY 1400 AGCTCTAACTAGATTAGATTGTTTGAACATCTACATCCACATTTCTTATGCACTGT 1459
Db 1315 AGCTCTAACTAGATTAGATTGTTTGAACATCTACATCCACATTTCTTATGCACTGT 1374
QY 1460 TCCCAATTTCTGTTCTACAAGCATGTTGTGGCAGAAACTGGAGACCGGCACTCTTA 1519
Db 1375 TCCCAATTTCTGTTCTACAAGCATGTTGTGGCAGAAACTGGAGACCGGCACTCTTA 1434

QY 1520 ATTTTACTTCAGCCATCGTATACCTCTTCTGACGTGATGGACCCCTCATACAAAGGTCCCT 1579
Db 1435 ATTTTACTTCAGCCATCGTATACCTCTTCTGACGTGATGGACCCCTCATACAAAGGTCCCT 1494
QY 1580 CTCATCATGTTCCAGTGAGAGGCGCGGATGCTTCTTCTCGGCATAGTAAACATTTTC 1639
Db 1495 CTCATCATGTTCCAGTGAGAGGCGCGGATGCTTCTTCTCGGCATAGTAAACATTTTC 1554
QY 1640 TTGGAACATATGTTTCACTTAATCACTACCAATATCTGGAACCTGTCTTACTCAGAC 1699
Db 1555 TTGGAACATATGTTTCACTTAATCACTACCAATATCTGGAACCTGTCTTACTCAGAC 1614
QY 1700 AGCACAGGTGTACAGAGCAGACAGATCTTCCAGATCAGAGGAGACCCCGGAG 1759
Db 1615 AGCACAGGTGTACAGAGCAGACAGATCTTCCAGATCAGAGGAGACCCCGGAG 1674
QY 1760 CTTCTGCTTCTCTACACTGGCATGCTGATGATCGTGACATGCCACATTCGCTTCTT 1819
Db 1675 CTTCTGCTTCTCTACACTGGCATGCTGATGATCGTGACATGCCACATTCGCTTCTT 1734
QY 1820 CCACATCTGTTGCACTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATCTTA 1879
Db 1735 CCACATCTGTTGCACTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATCTTA 1794
QY 1880 GTAGCAGGTGTTCTGCGAGGCTGTCTATGTGCTGCTGGCTGCCCAAGGACACTCCTG 1939
Db 1795 G-AGCAGGTGTTCTGCGAGGCTGTCTATGTGCTGCTGGCTGCCCAAGGACACTCCTG 1852
QY 1940 CAGAGCATTGTTGGTAAAGAACACTTAAAGAGGCACTTCAAGCAAGACATTCAGGCAAGATCCTCC 2059
Db 1853 CAGAGCATTGTTGGTAAAGAACACTTAAAGAGGCACTTCAAGCAAGACATTCAGGCAAGATCCTCC 1912
QY 2000 CAGAGCCTTTTGTAGAGGCTTCTGATGTCATTAAGACATTCAGGCAAGATCCTCC 2059
Db 1913 CAGAGCCTTTTGTAGAGGCTTCTGATGTCATTAAGACATTCAGGCAAGATCCTCC 1971
QY 2060 AACTGCAATATACCAACCTCTCTGAAATTTATTTGCTTATTTATATTTCTTTCTTT 2119
Db 1972 AACTGCAATATACCAACCTCTCTGAAATTTATTTGCTTATTTATATTTCTTTCTTT 2031
QY 2120 TTTTCTTAAAGATTTGCTCTGAAATGACATTTTCCATCTGAACTGGATGCAATC 2179
Db 2032 TTTTCTTAAAGTATGCTCTGAAATGACATTTTCCAT-TGAACCTGGATGCAATTC 2089
QY 2180 ATTTAGCCAACTCCAGTAAATTTTATTAATCTATACATAATATGTTTCTCAGCATA 2239
Db 2090 ATTTAGCCAACTCCAGTAAATTTTATTAATCTATACATAATATGTTTCTCAGCATA 2149
QY 2240 GGAGCTATGATTCAATTAATAAGTGGAGTCAAAACCGTAAATGCAATGTTTGTGTGT 2299
Db 2150 GGAGCTATGATTCAATTAATAAGTGGAGTCAAAACCGTAAATGCAATGTTTGTGTGT 2209
QY 2300 ATTTTCAATACAACTTAAATTTGCTTGTAAATAGTTCAGTGGATCTTGGAGTGG 2359
Db 2210 ATTTTCAATACAACTTAAATTTGCTTGTAAATAGTAC-AGTGGATCTTGGAGTGG 2268
QY 2360 GAATTTCTTGGTAAATTTATCTTGCATCTGAATGCTCATGATTACATATGAAATCGTTTG 2419
Db 2269 GAATTTCTTGGTAAATTTATCTTGCATCTGAATGCTCATGATTACATATGAAATCGTTTG 2328
QY 2420 ACATATCTTTAGACAGAAAAAGTAGTGGAGGGGAAATTTAGAGCTTGTGTGAC 2479
Db 2329 ACATATCTTTAGACAGAAAAAGTAGTGGAGGGGAAATTTAGAGCTTGTGTGAC 2387
QY 2480 TTTAGGAGTGTCT 2493
Db 2388 TTTAGGAGTGTCT 2401

RESULT 4

US-10-044-090-720

; Sequence 720, Application US/10044090


```
Db 180 CATGTCACCTCTTCTGACGATGAGCCGCTCATCAAAAGGTCCTCTCATCATGTTCC 239
QY 1593 AGTGAGAGCCAGGATGCTTCTTCTGCGCATAGTAACATTTTCTTGGAAATATGT 1652
Db 240 AGTGAGAGCCAGGATGCTTCTTCTGCGCATAGTAACATTTTCTTGGAAATATGT 299
QY 1653 TTCACTTAATCACTACCAATATCTGGAAGACCTGCTTCTACTCAGACAGCACCAGGTGA 1712
Db 300 TTCACTTAATCACTACCAATATCTGGAAGACCTGCTTCTACTCAGACAGCACCAGGTGA 359
QY 1713 CAGAAGCAGCAGACAAGATCTTCCAGATCAGCAGGAGAGCCCGGAGCCCTCTCTCTCC 1772
Db 360 CAGAAGCAGCAGACAAGATCTTCCAGATCAGCAGGAGAGCCCGGAGCCCTCTCTCTCC 419
QY 1773 TACATGCGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTTCCACATCTGGTTG 1832
Db 420 TACATGCGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTTCCACATCTGGTTG 479
QY 1833 CACTCGTCATGATGGGCTGCTGCATCTCCCTCAGTCCCAAAATCTAGTAGCCAGTGTT 1892
Db 480 CACTCGTCATGATGGGCTGCTGCATCTCCCTCAGTCCCAAAATCTAG - AGCCAAAGTGT 538
QY 1893 CTTGCAGAGGCTGCTCTATGCTGCTGCTGCCAAGGACACCTCCTGCAGAGCCATTTT 1952
Db 539 CTTGCAGAGGCTGCTCTATGCTGCTGCTGCCAAGGACACCTCCTGCAGAGCCATTTT 597
QY 1953 GGGTAAGGAACACTTCAAAAGAGGCAATGATCTTGTGCTGAGGCTCAGAGCCCTTTTG 2012
Db 598 GGGTAAGGAACACTTCAAAAGAGGCAATGATCTTGTGCTGAGGCTCAGAGCCCTTTTG 657
QY 2013 ATAGGCTTCTGATGCTCATTAAGACATTCAGGCAAGATGCTCCAACTGCAATATA 2072
Db 658 ATAGGCTTCTGATGCTCATTAAGACATTCAGGCAAGATGCTCCAACTGCAATATA 716
QY 2073 CCAACCTTCTCTGAAATATATTTGCTTATTTATATTTCTTTTCTTTTCTTAAAGAT 2132
Db 717 CCAACCTTCTCTGAAATATATTTGCTTATTTATATTTCTTTTCTTTTCTTAAAGAT - 775
QY 2133 TGCTCTGAATAGAAATGCAATTTTCACTGAACTGGAATGATCAATTTAGCCAAATCC 2192
Db 776 TGCTCTGAATAGAAATGCAATTTTCCAT - TGAACCTGGAATGATTTCAATTTAGCCAAATCC 834
QY 2193 AGTAATTTATTTATATATATATATATATATATATATATATATATATATATATATATAT 2252
Db 835 AGTAATTTATTTATATATATATATATATATATATATATATATATATATATATATATAT 894
QY 2253 ATTAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTAATTTTCAATACAC 2312
Db 895 ATTAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTAATTTTCAATACAC 954
QY 2313 AAACCTTAATTTGCTTGTGTAATTAAGTCAAGTGGATCTTGGAGTGGGATTTCTTGGTAA 2372
Db 955 AAACCTTAATTTGCTTGTGTAATTAAGTCAAGTGGATCTTGGAGTGGGATTTCTTGGTAA 1013
QY 2373 ATTATCTTGCATCTGAATGCTCATATATATATATATATATATATATATATATATATAT 2432
Db 1014 ATTATCTTGCATCTGAATGCTCATATATATATATATATATATATATATATATATATAT 1073
QY 2433 CAGAAAAAGTAGCTGAGTGAGGGGAAATTTATAGAGCTTGTGTGACTTTTAGGAGTAGC 2492
Db 1074 CAGAAAAAGTAGCTGAGTGAGGGGAAATTTATAGAGC-TGTGTGACTTTTAGGAGTAGC 1132
QY 2493 T 2493
Db 1133 T 1133
```

RESULT 8

US-09-867-701-2642
; Sequence 2642, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:

; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210321.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2642
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-867-701-2642

Query Match 11.8%; Score 303.8; DB 9; Length 507;
Best Local Similarity 97.3%; Pred. No. 1.1e-64;
Matches 330; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```
QY 2155 TTTCATCTGAACTGGATGCATATCAATTTAGCCAAATCCAGTAATTTTATATTAATCT 2214
Db 1 TTTCATCTGAACTGGATGCATTTTATTTAGCCAAATCCAGTAATTTTATATTAATCT 60
QY 2215 ATACATAATATGTTTCTCAGCATAGGAGCTATGATTCATTAAATTAAGTGGAGTCAAA 2274
Db 61 ATACATAATATGTTTCTCAGCATAGGAGCTATGATTCATTAAATTAAGTGGAGTCAAA 120
QY 2275 AGCTAAATGCAATGTTTGTGTAATTTTATTACACAACTTAATTTGTTGTTTAA 2334
Db 121 AGCTAAATGCAATGTTTGTGTAATTTTATTACACAACTTAATTTGTTGTTTAA 180
QY 2335 TAAGTCAAGTGGATCTTGGAGTGGGATTTCTTGGTAAATTAATTCGCACTTGAATGCT 2394
Db 181 TAAGTAC-AGTGGATCTTGGAGTGGGATTTCTTGGTAAATTAATTCGCACTTGAATGCT 239
QY 2395 CATGATTACATATGAATGCTTTGACATATCTTTAGACAGAAAAAGTAGCTGAG 2454
Db 240 CATGATTACATATGAATGCTTTGACATATCTTTAGACAGAAAAAGTAGCTGAG 299
QY 2455 GGGGAAATTTATAGAGCTTGTGTGACTTTAGGAGTAGCT 2493
Db 300 GGGGAAATTTATAGAGC-TGTGTGACTTTAGGAGTAGCT 337
```

RESULT 9

US-09-783-590-5564
; Sequence 5564, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5564
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (253)
; OTHER INFORMATION: n equals a,t,g, or c

```

; NAME/KEY: misc feature
; LOCATION: (261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (301)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (310)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (315)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (316)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (347)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (366)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (376)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (381)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (386)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (401)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (402)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (405)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (413)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (423)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (437)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (443)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (448)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature

; LOCATION: (453)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (460)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (464)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (467)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (470)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (479)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (485)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (487)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (488)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (492)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (493)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-5564

Query Match      8.5%; Score 219.4; DB 9; Length 497;
Best Local Similarity 90.7%; Pred. No. 9e-44;
Matches 255; Conservative 0; Mismatches 23; Indels 3; Gaps 2;

Qy      852 ATTCACTTAAGGCTGTCTCTTTAAAGTCACACCCCTCTTGATGGCCAGTGGAGGA 911
Db      2 ATTCCCTTAAGGCTGTCTCTTTAAAGTCACACCCCTCTTGATGGCCAGTGGAGGA 61

Qy      912 TCTTTGATGATGACCAACTATCATTAAGAGAGCTTTTACTCTTTGGCAATTATATTG 971
Db      62 TCTTTAATGATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGCAATTATATTG 121

Qy      972 TGGCTCCGAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTAGACAGGTACTGGC 1031
Db      122 TGGCTCCGAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTAGACAGGTACTGGC 181

Qy      1032 ACAGCAGGGCCACTATGCCATAGCCACTG-AGTGTAGTGGGGTCTCCAGCAGCTGCC 1090
Db      182 ACAGCAGGGCCACTATGCCATAGCCACTGAGGTGAGTGGGGGTCTCCAGCAGCTGCC 241

Qy      1091 CCAGCCCATACCCCTGTGTA--CCAGTGGCCAGAGGACCTG 1129
Db      242 CCAGCCCATCANCTGTGTNACCAGTGGGGCCAGAGGACCTG 282

RESULT 10
US-10-240-425-827/c
; Sequence 827, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherif, Uwe
```

```
/ APPLICANT: Vockley, Joseph G.
/ TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
/ FILE REFERENCE: 44921-5026
/ CURRENT APPLICATION NUMBER: US/10/240,425
/ PRIOR FILING DATE: 2002-09-30
/ PRIOR APPLICATION NUMBER: PCT/US01/09847
/ PRIOR FILING DATE: 2001-03-28
/ PRIOR APPLICATION NUMBER: US 60/193,446
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 1588
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 827
/ LENGTH: 603
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: n equals a,t,g, or c
/ US-10-240-425-827

Query Match      5.8%; Score 149.8; DB 16; Length 603;
Best Local Similarity 97.7%; Pred. No. 1.7e-26;
Matches 173; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 2317 TTAATTGCTCTGTTAAATAAGTTCAGTGGATCTTCGAGTGGGATTTCTTGGTAAATTA 2376
Db 603 TTAATTGCTCTGTTAAATAAGTAC-AGTGAATCTTGGAGTGGATTTCTTGGTAAATTA 545

QY 2377 TCTTGACATTAAGTCTCATGATATACATATGAATGAAATCGCTTTGACATATCTTAGACAGA 2436
Db 544 TCTTGACATTAAGTCTCATGATATACATATGAATGAAATCGCTTTGACATATCTTAGACAGA 485

QY 2437 AAAAAGTAGCTGAGTGGGGGAAATTTATAGAGCTTGATGACTTTAGGGAGTAGCT 2493
Db 484 AAAAAGTAGCTGAGTGGGGGAAATTTATAGAGC-TGTGTGACTTTAGGGAGTAGGT 429

RESULT 11
US-09-783-590-11496
/ Sequence 11496, Application US/09783590
/ Patent No. US20020110850A1
/ GENERAL INFORMATION:
/ APPLICANT: Dillon, Patrick J.
/ APPLICANT: Haseltine, William A.
/ APPLICANT: Li, Haodong
/ APPLICANT: Rosen, Craig A.
/ APPLICANT: Ruben, Steven M.
/ TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
/ FILE REFERENCE: PO-16.2C1
/ CURRENT APPLICATION NUMBER: US/09/783,590
/ PRIOR FILING DATE: 2000-02-15
/ PRIOR APPLICATION NUMBER: 08/420,856
/ PRIOR FILING DATE: 1995-04-12
/ PRIOR APPLICATION NUMBER: 08/346,731
/ PRIOR FILING DATE: 1994-11-21
/ NUMBER OF SEQ ID NOS: 12485
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 11496
/ LENGTH: 390
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (29)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (60)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (71)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (151)
/ OTHER INFORMATION: n equals a,t,g, or c

/ APPLICANT: Vockley, Joseph G.
/ TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
/ FILE REFERENCE: 44921-5026
/ CURRENT APPLICATION NUMBER: US/10/240,425
/ PRIOR FILING DATE: 2002-09-30
/ PRIOR APPLICATION NUMBER: PCT/US01/09847
/ PRIOR FILING DATE: 2001-03-28
/ PRIOR APPLICATION NUMBER: US 60/193,446
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 1588
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 827
/ LENGTH: 603
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: n equals a,t,g, or c
/ US-10-240-425-827

Query Match      4.9%; Score 126.2; DB 9; Length 390;
Best Local Similarity 80.5%; Pred. No. 9.6e-21;
Matches 182; Conservative 0; Mismatches 38; Indels 6; Gaps 3;

QY 1563 TCATCAAAAGTCCCTCTCATCATGTTCCAGTGAGAGCCAGGATTCCTTCTTCCTG 1622
Db 10 TCATCAAAAGTCCCTCTCATCATGTTCCAGTGAGAGCCAGGATTCCTTCTTCCTG 69

QY 1623 GCATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAATATCTGGAAG 1682
Db 70 GNATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAATCACTACCAATATCTGGAAG 129

QY 1683 -ACCTCTCTTACTCAGA--CAGCACAGGTGTACAGAAG--CAGCAGACAAAGATCTTCC 1736
Db 130 AACCTCTCTTACTCAGAACAGNCACCAGGTGTACAGAAGGCAGCAGACAAAGATCTTCC 189

QY 1737 AGATCAGAGGGAGACCCCGAGCCTCTGCTTCTCTTACTACTGGCA 1782
Db 190 AGATCAGAGGGAGACCCCGAGCCTCTGATTTCTCCCTNACA 235

RESULT 12
US-10-156-761-7362
/ Sequence 7362, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
```

APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.1%; Score 79.8; DB 15; Length 9025608;
Best Local Similarity 45.5%; Pred. No. 5.6e-07;
Matches 322; Conservative 0; Mismatches 382; Indels 3; Gaps 1;

QY 1 CGGGGCGGAGGGCGGCTGCTGAGGCGCCATGGGCTTCGCGCGCGCGGCTTCCTGGCGCG 60
Db 8812441 CGGCGCTCGGCTCGCGAACCAGGAGACCCACCGCTGCTCCAGAACCGCGCGT 8812500
QY 61 GCACACTGTGGGCGCGCTGCTGGGGCGGCGGGGTGTGGCTGGGGCCATGCTCGCC 120
Db 8812501 ACCGCACCGCGGTGGTTCGTGCGCCCTGGTGTGCTGCGCGAGACCTGGAGAGCGCGGC 8812560
QY 121 GTGCGGCTTGTGCTGGAGCTTCGCGACTGCACCTGGCTCACTTCGCCCTTAGCGCGCGA 180
Db 8812561 GSGCGCGCGCGCGTTCGCGGCTCGACTACGCGGCGGAGACACGACGTGACGTGACGA 8812620
QY 181 CGGCGCGCGGCGAGCAGACGCGCGCGCGCGCGCGCTGGCGCGCGCTGTGGGCGCGCGG 240
Db 8812621 GACCCACCGCGAGGCTTACACGCGCGGACACGCGGAGGATATCCGGCACACCGCGA 8812680
QY 241 GCGGAGTACTGTGCTGTGCTGGCGCGGTGACCGCGGAGCGCGGTGCGGGCGCGCGGTCG 300
Db 8812681 ACACGGGAGCGCGAGTGGCGGTTCACCTGCGCGCGCGGTTCGCGTGCACGCTGCG 8812740
QY 301 GGGCGCGCGCTGCACAGCGCGCTGCTGACCGCGCGCGCGCGCGCGCGCTTCAGCGCGGT 360
Db 8812741 GCTGCGCGCGCTGCACAAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 8812800
QY 361 CCAGCTGTCTAGGCTGTCTGTCTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 8812801 CGGCAGACTGTGTGTGACGACTCCAGTCAGGCGCGCGCGCGCGCGCGCGCGCGCG 8812860
QY 421 CTTCTGTGTGGCGGACCG 480
Db 8812861 GGGCGCTGTTCAGCTGCG 8812920
QY 481 GGGCGCGCTGCAGAGGCG 540
Db 8812921 CTTGCGCTGCAAGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8812979
QY 541 CCAGCTGTGGCAGCGCGCTCTGGGAGGTGCAAGAAGGAGGCGCGCGCGCGCGCGCGCG 600
Db 8812980 --AGACGGCG 8813037
QY 601 ACAGGTGTGCGCGTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 8813038 GCACCG 8813097
QY 661 CGTGTCTTCCGGACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 707
Db 8813098 CGCGTCTGTGACGAGGTCAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8813144

NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 7362
LENGTH: 2118
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2118)
US-10-156-761-7362

Query Match 3.1%; Score 79.8; DB 15; Length 2118;
Best Local Similarity 45.5%; Pred. No. 7.2e-09;
Matches 322; Conservative 0; Mismatches 382; Indels 3; Gaps 1;

QY 1 CGGGGCGGAGGGCGGCTGCTGAGGCGCCATGGGCTTCGCGCGCGCGGCTTCGCGCGG 60
Db 219 CGGCGCTCGGCTCGCGAACCAGGAGACCCACCGCTGCTCCAGAACCGCGCGT 278
QY 61 GCACACTGTGGGCGCGCTGCTGGGCGGCGCGGGGTGTGGCTGGGGCCATGCTCGCC 120
Db 279 ACCGCACCGCGGTGGTTCGTGCGCCCTGGTGTGCTGCGCGAGACCTGGAGAGCGCGGC 338
QY 121 GTGCGGCTTGTGCTGGAGCTTCGCGACTGCACCTGGCTCACTTCGCCCTTAGCGCGCGA 180
Db 339 GGGCGCGCGCGCGCTCGCGGTTCGACTACGCGGCGGAGACACGACGTGACGTGACCGA 398
QY 181 CGGCGCGCGGCGAGCAGACGCGCGCGCGCGCGCGCGCGCTGGCGCGCGCTGTGGGCGCGCG 240
Db 399 GACCCACCGCGAGGCTTACACGCGGAGACCGCGGAGGATATCCGGCACACCGCGA 458
QY 241 GCGGAGTACTGTGTGCTGTGCGCGGTGACCGCGGAGCGCGGTGCGGGCGCGGGTCCG 300
Db 459 ACACGGGAGCGCGAGTGGCGGTTCACCTGCGGCGCGGTTCGCGTGCACGTGCGCTACCG 518
QY 301 GSGCGCGCGCTGCACAGCGCGCTGTGACCGCGCGCGCGCGCGCGCGCTTCAGCGCGGT 360
Db 519 GCTGCGCGCGCTGCACACACACCCATGAGCGCGCACCGCGCGCGCGCTTCGCGAGG 578
QY 361 CCAGCTGTCTAGGCTGTCTGTCTACTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 579 CGGCAGACTGTGTGACGACTCCAGTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 638
QY 421 CTTCTGTGCGGCG 480
Db 639 GGGCGCTGTTCAGCTGCGCGGAGGACCGGTTCACCGTGTGCGCGCGCGCGCGCGCGCG 698
QY 481 GGGCGCTGTGCAGGAGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 699 CTTGCGCTGCAAGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 757
QY 541 CCAGCTGTGGCAGCGCGCTCTGGGAGGTGCAAGAAGGAGGCGCGCGCGCGCGCGCGCG 600
Db 758 --AGACGGCG 815
QY 601 ACAGGTGTGCGCGTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 816 GCACCG 875
QY 661 CGTGTCTTCCGGACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 707
Db 876 CGCGTCTGTGACGAGGTCAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 922

RESULT 13
US-10-156-761-1
Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI

This Page Blank (uspto)

361	DB	CGAGCTGCTCAGGGTGTCTGTCTACTGCGCGGGGGCCAGCGCGGCCACAGCAAGG	420
421	QY	CTTCCTGCTGCGGACCCCTGGATGACCTGACACCCGGCAAGCGTGTCTGAGCTGCT	480
421	DB	CTTCCTGCTGCGGACCCCTGGATGACCTGACACCCGGCAAGCGTGTCTGAGCTGCT	480
481	QY	GGGGCCCTGCCAGSAGSACCAAGCCCGGCACTTTGGGGGAGTTGAGSACCACCGCGCGG	540
481	DB	GGGGCCCTGCCAGSAGSACCAAGCCCGGCACTTTGGGGGAGTTGAGSACCACCGCGCGG	540
541	QY	CCAGCTGTGGCAGGCGCTCTGGGAGGTGCAAGACGGCAGGCGGCTGCAGGTGGGCTGCGC	600
541	DB	CCAGCTGTGGCAGGCGCTCTGGGAGGTGCAAGACGGCAGGCGGCTGCAGGTGGGCTGCGC	600
601	QY	ACAGGTGTGCGCGCTCCCGAGCGCCCGCTGCGACCCCGGTGGTGCAGACTTGCACGATTG	660
601	DB	ACAGGTGTGCGCGCTCCCGAGCGCCCGCTGCGACCCCGGTGGTGCAGACTTGCACGATTG	660
661	QY	CGTGGTCTTCCCGACCGGGAAGCGCGCGGCGGTTTGGAGAGGTGATACCTCTTTAT	720
661	DB	CGTGGTCTTCCCGACCGGGAAGCGCGCGGCGGTTTGGAGAGGTGATACCTCTTTAT	720
721	QY	TCCTGAAGCCGGGCGAGTGTGTGACCTGGTGCACCGAGTGGCCCAAAACAGATCCAGAAAG	780
721	DB	TCCTGAAGCCGGGCGAGTGTGTGACCTGGTGCACCGAGTGGCCCAAAACAGATCCAGAAAG	780
781	QY	AAAGTTCCAGGTTGTTGCCATCGSAGACTGGATGCCACGGGTAAACACAGGTGACCCA	840
781	DB	AAAGTTCCAGGTTGTTGCCATCGSAGACTGGATGCCACGGGTAAACACAGGTGACCCA	840
841	QY	GTCAGTGGCAGATTCACTTAAGGTGTCTCTTAAAGTCAACACCTCTTGCAATGGCCCA	900
841	DB	GTCAGTGGCAGATTCACTTAAGGTGTCTCTTAAAGTCAACACCTCTTGCAATGGCCCA	900
901	QY	GTGAGAGAAATCTTTGATGATGAACCAACTATCATTTAGAAGAGCTTTTACTCTTTGGG	960
901	DB	GTGAGAGAAATCTTTGATGATGAACCAACTATCATTTAGAAGAGCTTTTACTCTTTGGG	960
961	QY	CAATTAATTTGGCCCTCCGAAATAGCTAAAGATCTGCCAAATCTCTGTGATGTAGA	1020
961	DB	CAATTAATTTGGCCCTCCGAAATAGCTAAAGATCTGCCAAATCTCTGTGATGTAGA	1020
1021	QY	CAGGTACTGGCAGCAGCGCCACTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCA	1080
1021	DB	CAGGTACTGGCAGCAGCGCCACTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCA	1080
1081	QY	GCACCTGCCCGCCCATCACCTGTGTACAGCTGGCCAGAGACCTGCTCAAACTGA	1140
1081	DB	GCACCTGCCCGCCCATCACCTGTGTACAGCTGGCCAGAGACCTGCTCAAACTGA	1140
1141	QY	CTTTATCTGTGCTCACTGTGAGTCTCTGAGGAGAGGTGTCAGAGGCTGCAGGGCCGGGG	1200
1141	DB	CTTTATCTGTGCTCACTGTGAGTCTCTGAGGAGAGGTGTCAGAGGCTGCAGGGCCGGGG	1200
1201	QY	CATGGAGAGNACAGGAGAGAGCAACTTTCAGGCGCAACAGTGTGTTTGTGTAAGAGGT	1260
1201	DB	CATGGAGAGNACAGGAGAGAGCAACTTTCAGGCGCAACAGTGTGTTTGTGTAAGAGGT	1260
1261	QY	AGAAATGCTTACCAGCGGATGGAGATCTCTGGCTGCCATGTGGTTCATGCCAGCCCTC	1320
1261	DB	AGAAATGCTTACCAGCGGATGGAGATCTCTGGCTGCCATGTGGTTCATGCCAGCCCTC	1320
1321	QY	CAGAGAAAGGTCTCGACAGCATTAAGCCTTAATCCAGAAATGTTTTAGTGAACCGTA	1380
1321	DB	CAGAGAAAGGTCTCGACAGCATTAAGCCTTAATCCAGAAATGTTTTAGTGAACCGTA	1380
1381	QY	GTTACTCTGGCAGGTGCCAGTCTAACATGATTAAGTGTGTTTGAACCACTTACATCC	1440
1381	DB	GTTACTCTGGCAGGTGCCAGTCTAACATGATTAAGTGTGTTTGAACCACTTACATCC	1440
1441	QY	ACCAATTTGTATGCACTGTCTCCAAATTTCTGTTCTAACAGCACTGTTGTGGCAGAAA	1500
1441	DB	ACCAATTTGTATGCACTGTCTCCAAATTTCTGTTCTAACAGCACTGTTGTGGCAGAAA	1500
1501	QY	CTGGAGACAGGCACTTAAATTTTACTTTAGCCATCGTACCCTCTTCTGACTGATGGACC	1560
1501	DB	CTGGAGACAGGCACTTAAATTTTACTTTAGCCATCGTACCCTCTTCTGACTGATGGACC	1560
1561	QY	CGTCATCACAAAGTCCCTCTCATCATGTTTCCAGTGAGAGGCCAGGATTTGCTTTCTTCC	1620
1561	DB	CGTCATCACAAAGTCCCTCTCATCATGTTTCCAGTGAGAGGCCAGGATTTGCTTTCTTCC	1620
1621	QY	TGGCATAGTAAACATTTTCTTGGAACATATGTTTCACTTAATCATACCAATATCTGGA	1680
1621	DB	TGGCATAGTAAACATTTTCTTGGAACATATGTTTCACTTAATCATACCAATATCTGGA	1680
1681	QY	AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAGCAGACAGACAAAGATCTTCCAGAT	1740
1681	DB	AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAGCAGACAGACAAAGATCTTCCAGAT	1740
1741	QY	CAGCAGGAGAGACCCCGAGCCTCTGCTTCTCCTACACTGGCATGTGATGAGATCGTAGC	1800
1741	DB	CAGCAGGAGAGACCCCGAGCCTCTGCTTCTCCTACACTGGCATGTGATGAGATCGTAGC	1800
1801	QY	ATGCCCCACATTTGGCTTCTTCCACATCTGGTTCGATCTGATGGGCTCGCTGCATCT	1860
1801	DB	ATGCCCCACATTTGGCTTCTTCCACATCTGGTTCGATCTGATGGGCTCGCTGCATCT	1860
1861	QY	CCCTCAGTCCCAAATTTAGTAGCCAAAGTGTTCCTGCAGAGGCTGTCTATGTCTCTGGC	1920
1861	DB	CCCTCAGTCCCAAATTTAGTAGCCAAAGTGTTCCTGCAGAGGCTGTCTATGTCTCTGGC	1920
1921	QY	TGCCAAAGGAGACACTCCTGCAGAGCCATTTTGGGTAAAGAACACTTACAAGAGGCAT	1980
1921	DB	TGCCAAAGGAGACACTCCTGCAGAGCCATTTTGGGTAAAGAACACTTACAAGAGGCAT	1980
1981	QY	TGATCTTGTGTGAGGCTCAGAGCCCTTTTGTATAGGCTTCTGATGCTCATTCATAAAGAC	2040
1981	DB	TGATCTTGTGTGAGGCTCAGAGCCCTTTTGTATAGGCTTCTGATGCTCATTCATAAAGAC	2040
2041	QY	ATTCAAGCCAGAGTGTCCAACTGCAAAATATACCAACCTTCTCTGAATTAATATTTGCTT	2100
2041	DB	ATTCAAGCCAGAGTGTCCAACTGCAAAATATACCAACCTTCTCTGAATTAATATTTGCTT	2100
2101	QY	ATTATATTTCTTTCTTTTCTTAAAGATTTGGCTCTGAAATAGAGTGCACATTTTCCA	2160
2101	DB	ATTATATTTCTTTCTTTTCTTAAAGATTTGGCTCTGAAATAGAGTGCACATTTTCCA	2160
2161	QY	TCTGAACCTGATGCATATCATTTAGCCAACTCCAGTAATTTAATTAATTAATCTATACAT	2220
2161	DB	TCTGAACCTGATGCATATCATTTAGCCAACTCCAGTAATTTAATTAATTAATCTATACAT	2220
2221	QY	AATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTAAGAGTGGAGTCAAAACGCTA	2280
2221	DB	AATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTAAGAGTGGAGTCAAAACGCTA	2280
2281	QY	AATGCAATGTTTGTGTATTTTCAATTAACAACTTAATTTGTCTTGTAAATAAGTT	2340
2281	DB	AATGCAATG	


```
RESULT 2
US-10-277-032-1
; Sequence 1, Application US/10277032
; Patent No. 6664087
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305 DIV
; CURRENT APPLICATION NUMBER: US/10/277,032
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: HomoSapien
US-10-277-032-1

Query Match      100.0%; Score 2571; DB 4; Length 2571;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCCGAGGGCGGCTCGCTGAGCGGCCCATGGCTTCGGCCGCGGCTCGCTCGCGG 60
DB 1 CGGGCCGAGGGCGGCTCGCTGAGCGGCCCATGGCTTCGGCCGCGGCTCGCTCGCGG 60

QY 61 GCCACTGTGGGGCCGCTGCTGGGGGGGGGGGCTGGGGTGGGGGCAATGGCTCGGC 120
DB 61 GCCACTGTGGGGCCGCTGCTGGGGGGGGGGGCTGGGGTGGGGGCAATGGCTCGGC 120

QY 121 GTGCCGCTTCTGCTGGAGCTTCCCGACTGCACTGGCTGGCTCACTTCGCTAGGCGCCGA 180
DB 121 GTGCCGCTTCTGCTGGAGCTTCCCGACTGCACTGGCTGGCTCACTTCGCTAGGCGCCGA 180

QY 181 CGCCCCCGCGAGCGAGCGCCCGCCCGCCCGCTGGGGGCGCTGTGGGGCCCGCGGA 240
DB 181 CGCCCCCGCGAGCGAGCGCCCGCCCGCCCGCTGGGGGCGCTGTGGGGCCCGCGGA 240

QY 241 GCGCAGCTACTGCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
DB 241 GCGCAGCTACTGCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300

QY 301 GCGCGCGCGGCTGCAACGCGCTGCTGCAACGCGCTGCGCGCGCGCGCTTCCAGCGGTG 360
DB 301 GCGCGCGCGGCTGCAACGCGCTGCTGCAACGCGCTGCGCGCGCGCGCTTCCAGCGGTG 360

QY 361 CCAGCTGCTCAGGCTGCTGTGCTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CCAGCTGCTCAGGCTGCTGTGCTACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 420

QY 421 CTTCTGCTGCGGACCCCGCTGATGACCTGACCCCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTTCTGCTGCGGACCCCGCTGATGACCTGACCCCGCGCGCGCGCGCGCGCGCG 480

QY 481 GGGCGGCTGCGCAGGAGGACCAACGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCG 540
DB 481 GGGCGGCTGCGCAGGAGGACCAACGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCG 540

QY 541 CCAGCTGTGGCAGCGCTCTGGAGAGTGCAAGACGCGCGCGCGCGCGCGCGCGCG 600
DB 541 CCAGCTGTGGCAGCGCTCTGGAGAGTGCAAGACGCGCGCGCGCGCGCGCGCGCG 600

QY 601 ACAGTTCGTCGCGGAGCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 601 ACAGTTCGTCGCGGAGCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

QY 661 CGTGGTCTTCCGCGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 CGTGGTCTTCCGCGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
```

QY	1801	ATGCCACATGGCTTCTTCCACATCTGGTTGCACTCGTCATGATGGGTCGCTGCATCT	1860
DB	1801	ATGCCACATGGCTTCTTCCACATCTGGTTGCACTCGTCATGATGGGTCGCTGCATCT	1860
QY	1861	CCCTCAGTCCCAAAATCTAGTAGCCAAAGTGTTCCTGCAGAGGCTGTCTATGTGCTCGC	1920
DB	1861	CCCTCAGTCCCAAAATCTAGTAGCCAAAGTGTTCCTGCAGAGGCTGTCTATGTGCTCGC	1920
QY	1921	TGCCAAGGACACACTCTGTCAGAGCCATTTTGGGTAAAGAACACATTACAAAGAAGCAT	1980
DB	1921	TGCCAAGGACACACTCTGTCAGAGCCATTTTGGGTAAAGAACACATTACAAAGAAGCAT	1980
QY	1981	TGATCTTGCTCTGAGGCTCAGAGCCCTTTTGATAGGCTTCTGATTCGATTCATAAAGAC	2040
DB	1981	TGATCTTGCTCTGAGGCTCAGAGCCCTTTTGATAGGCTTCTGATTCGATTCATAAAGAC	2040
QY	2041	ATTCAAGCCAAAGATGTCCTCAACTGCAAATATACCAACCTTCTCTGAATATATTTTGCTT	2100
DB	2041	ATTCAAGCCAAAGATGTCCTCAACTGCAAATATACCAACCTTCTCTGAATATATTTTGCTT	2100
QY	2101	ATTATATATTTCTTTTCTTTTCTTAAGAAATTTGGCTCTGAAATAGAAATGCACATTTTCCA	2160
DB	2101	ATTATATATTTCTTTTCTTTTCTTAAGAAATTTGGCTCTGAAATAGAAATGCACATTTTCCA	2160
QY	2161	TCTGAACCTGGATGCATATCATTTAGCCAAATCCAGTAATTTATTATATTAATCTATACAT	2220
DB	2161	TCTGAACCTGGATGCATATCATTTAGCCAAATCCAGTAATTTATTATATTAATCTATACAT	2220
QY	2221	AATAATGTTTCTCAGCATAGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAACGCTA	2280
DB	2221	AATAATGTTTCTCAGCATAGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAACGCTA	2280
QY	2281	AATGCATGTTTGTTGTGTATTTTCATTACACAACTTAATTTCTCTTTGTATAAATAGTT	2340
DB	2281	AATGCATGTTTGTTGTGTATTTTCATTACACAACTTAATTTCTCTTTGTATAAATAGTT	2340
QY	2341	CAAGTGGATCTTGGAGTGGGATTTCTTTGGTAAATTTATCTTGCACTTGAATGTCATGAT	2400
DB	2341	CAAGTGGATCTTGGAGTGGGATTTCTTTGGTAAATTTATCTTGCACTTGAATGTCATGAT	2400
QY	2401	TACATATGAATTCGCTTTTGACATATCTTTAGACAGAAAAGTAGTCTGAGTGGGGGAA	2460
DB	2401	TACATATGAATTCGCTTTTGACATATCTTTAGACAGAAAAGTAGTCTGAGTGGGGGAA	2460
QY	2461	ATTATAGAGCTTTGTGACTTTTAGGAGTAGCTGTCTCTTATACACATACTCAAGCCCTG	2520
DB	2461	ATTATAGAGCTTTGTGACTTTTAGGAGTAGCTGTCTCTTATACACATACTCAAGCCCTG	2520
QY	2521	AAGCCTTGCACTGCTCTGCAGCGTCCGACTAAAGAGGGGGCTTTTTCACCC	2571
DB	2521	AAGCCTTGCACTGCTCTGCAGCGTCCGACTAAAGAGGGGGCTTTTTCACCC	2571

RESULT 3

```

US-09-984-880-3
; Sequence 3, Application US/09984880
; Patent No. 6489153
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001305
; CURRENT APPLICATION NUMBER: US/09/984,880
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 20966
; TYPE: DNA
; ORGANISM: HomoSapien
US-09-984-880-3

```

```
QY 2262 AAGTGAGTCAAAACGTAATGCAATGTTGTTGTTGTTATTTTCATTACACAACTTAAT 2321
Db 18739 AAGTGAGTCAAAACGTAATGCAATGTTGTTGTTGTTATTTTCATTACACAACTTAAT 18798
QY 2322 TTGCTTTGTTAAATAAGTCAAGTGGATCTTGAGTGGGATTTCTTGGTAAATATCTTG 2381
Db 18799 TTGCTTTGTTAAATAAGTAC-AGTGGATCTTGAGTGGGATTTCTTGGTAAATATCTTG 18857
QY 2382 CACTGAAATGCTCATGATTACATATGAATATGAAATCGCTTTGACATATCTTTAGACAGAAAAA 2441
Db 18858 CACTGAAATGCTCATGATTACATATGAATATGAAATCGCTTTGACATATCTTTAGACAGAAAAA 18917
QY 2442 GTAGCTGAGTGGGGGAAATTATAGAGCTTGTGCTTGGAGTAGCT 2493
Db 18918 GTAGCTGAGTGGGGGAAATTATAGAGC-TGTGTGACTTTAGGAGTAGGT 18968

RESULT 4
US-10-277-032-3
; Sequence 3, Application US/10277032
; Patent No. 6664087
; GENERAL INFORMATION:
; APPLICANT: Ming-Hui WEI
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001305 DIV
; CURRENT APPLICATION NUMBER: US/10/277,032
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/984,880
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 20966
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-277-032-3

Query Match 44.0%; Score 1130.4; DB 4; Length 20966;
Best Local Similarity 98.1%; Pred. No. 9.3e-260;
Matches 1228; Conservative 0; Mismatches 16; Indels 8; Gaps 8;

QY 1243 TGTGTTTCTCAAAAGGTAGAAATGTCCTACAGCGGATGGAGATCTCGGCTGCCATGT 1302
Db 17724 TTGCGCTGTCCCGAGGTAGAAATGTCCTACAGCGGATGGAGATCTCGGCTGCCATGT 17783
QY 1303 GGTGTGATGCCAGCCCTCCAGAGAAAAGTCTCGCAGACAGTATTAAAGCCTAATCCAGAA 1362
Db 17784 GGTGTGATGCCAGCCCTCCAGAGAAAAGTCTCGCAGACGCTAATTAAAGCCTAATCCAGAA 17843
QY 1363 TAGTTTATGTAACCGTAGTTACTCTGGCCAGGTGCCAGTCTPACTAGATTAGATGTTG 1422
Db 17844 TAGTTTATGTAACCGTAGTTACTCTGGCCAGGTGCCAGTCTPACTAGATTAGATGTTG 17903
QY 1423 TTGAAACATCTACATCCACCATTTGTTATGAGTGTCCCAATTTCTGTTCTACAAGC 1482
Db 17904 TTGAAACATCTACATCCACCATTTGTTATGAGTGTCCCAATTTCTGTTCTACAAGC 17963
QY 1483 ATGTTCTGTGGCGAAGAACTGGAGACAGGCATCTTAAATTTTACTTCAGCCATCGTACCC 1542
Db 17964 ATGTTCTGTGGCGAAGAACTGGAGACAGGCATCTTAAATTTACTTCAGCCATCGTACCC 18023
QY 1543 TC-TTCTGACTGATGAGACCGGTATCACAAGGTCCCTCTCATCATGTTTCCAGTGAAGG 1601
Db 18024 TC-TTCTGACTGATGAGACCGGTATCACAAGGTCCCTCTCATCATGTTTCCAGTGAAGG 18083
QY 1602 CCAGCCATGCTTTCTTCTCGCATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAA 1661
Db 18084 CCAGCCATGCTTTCTTCTCGCATAGTAAACATTTTCTTGGAAACATATGTTTCACTTAA 18143
QY 1662 TCACTACCAAAATCTGGAAGACCTGCTTCTACTCAGACAGCACCGGTGTACAGAGCAG 1721
```

```
Db 18144 TCACTACCAAAATATCTGGAAGACCTGCTTCTACTCAGACAGCACCGGTGTACAGAGCAG 18203
QY 1722 CAGACAAGATCTTCCAGATCAGCAGGGAGACCCCGAGGCTCTGCTTCTCCTACACTGGC 1781
Db 18204 CAGACAAGATCTTCCAGATCAGCAGGGAGACCCCGAGGCTCTGCTTCTCCTACACTGGC 18263
QY 1782 ATGCTCATGAGATCTGATGATGCCACATTTGGCTTCTTCCACATCTGGTGTGACTCGTCA 1841
Db 18264 ATGCTCATGAGATCTGATGATGCCACATTTGGCTTCTTCCACATCTGGTGTGACTCGTCA 18323
QY 1842 TGATGGCTCGCTGATCTCCCTCAGTCCAAATTTCTAGTAGCAAGTGTCTCTGCAGAG 1901
Db 18324 TGATGGCTCGCTGATCTCCCTCAGTCCAAATTTCTAG-AGCAAGTGTCTCTGCAGAG 19382
QY 1902 GCTGTCTATGTGTCTGCTGCTGCCAAGGACACTCTCTGCAGAGCCATTTTGGGTAAAGA 1961
Db 18383 GCTGTCTATGTGTCTGCTGCTGCCA-AGCACACTCTCTGCAGAGCCATTTTGGGTAAAGA 18441
QY 1962 ACATTTACAAAGAGGATTTGATCTTGTCTCTGAGGCTCAGAGCCCTTTTGATAGGCTTC 2021
Db 18442 ACATTTACAAAGAGGATTTGATCTTGTCTCTGAGGCTCAGAGCCCTTTTGATAGGCTTC 18501
QY 2022 TGATGTCAATTCATAAAGACATTCAGCCCAAGATGCTCCAACTGCAATATACCAACTTC 2081
Db 18502 TGA-GTCATATATAAGACATTCAGCCCAAGATGCTCCAACTGCAATATACCAACTTC 18560
QY 2082 TCTGAATATATTTTGGCTTATTTATATTTCTTTCTTTTCTTTTCTTAAAGAAATGGCTCTGA 2141
Db 18561 TCTGAATATATTTTGGCTTATTTATATTTCTTTCTTTTCTTTTCTTAAAGTA-TGGCTCTGA 18619
QY 2142 ATAGAATGCACATTTTCCATCTGAACTGCATGCATATCATTTAGCCAAATCCAGTAATTA 2201
Db 18620 ATAGAATGCACATTTTCCAT-TGAACCTGCATGCATTTAGCCAAATCCAGTAATTTA 18678
QY 2202 TTTATATTAATCTATACATAATATGTTTCTCTCAGCATAGGAGTATGATTTCATTAAATTA 2261
Db 18679 TTTATATTAATCTATACATAATATGTTTCTCTCAGCATAGGAGTATGATTTCATTAAATTA 18738
QY 2262 AAGTGAGTCAAAACGCTAAATGCAATGTTGTTGTTGTTTTCATTACACAACTTAAT 2321
Db 18739 AAGTGAGTCAAAACGCTAAATGCAATGTTGTTGTTGTTTTCATTACACAACTTAAT 18798
QY 2322 TTGCTCTTGTAAATAAGTTCAGTGGATCTTGAGTGGGATTTCTTGGTAAATATCTTG 2381
Db 18799 TTGCTCTTGTAAATAAGTAC-AGTGGATCTTGAGTGGGATTTCTTGGTAAATATCTTG 18857
QY 2382 CACTTGAATGCTCATGATTACATATGAATCGCTTTCATATCTTTAGACAGAAAAA 2441
Db 18858 CACTTGAATGCTCATGATTACATATGAATCGCTTTCATATCTTTAGACAGAAAAA 18917
QY 2442 GTAGCTGAGTGGGGGAAATTATAGAGCTTGTGACTTTAGGAGTAGGT 2493
Db 18918 GTAGCTGAGTGGGGGAAATTATAGAGC-TGTGTGACTTTAGGAGTAGGT 18968

RESULT 5
US-09-266-965-51
; Sequence 51, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456US1
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
```

Query Match 3.0%; Score 76; DB 4; Length 53500;
 Best Local Similarity 50.8%; Pred. No. 1.6e-07;
 Matches 181; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
 QY 235 CCCGAGGCGACACTACTCGTGTGGTGGTCCCGTACCCCGACGCCGGCTCGGGGGCCG 294

Qy	35	GCCTTCGCCCGCCGCGTCTCTCGCGGGGACATGTCGGGGCGCTGCTGGCGCGCGCGG	94
Db	1205	GCTGCGCGCGGTGCGGCGCGCTCGGGCTCTGCCCGGGGAGCGCAAGCGTCTCGAG	1146
Qy	95	GTCTGCGTGGGGGCATGGTCCCGCGCTTCGTCTCTGGAGTTCGCGATGTGACC	154
Db	1145	ACGGTGGCGTCCAGGCGGCGCGCGCGCTCTCTTCCAGCAGACCGCTCACGCG	1086
Qy	155	CTGGCTCATCTTCGCCTTAGCGCGGACGCGCCCGCGGACGACAGCGCCCCCGACCGCCG	214
Db	1095	CCGCGCGCTCAGGCGCGCGCTGCGCTGCTGTACCTCCCGAGGTAGCGCAGCATTCGG	1020

```
QY 215 CTGGCGGCGCTGTTGGGGCCCGGAGCGAGCTACTCGCTGTGCTGCGTGCACCCCG 274
Db 1025 CAGGCT-|||||
QY 275 GACCGCGGCTGGGGCCCGGGTCCGGGCGGCGCGCTCACACAGCGCTGTGTGACACAG 334
Db 979 TCGGGCGGCGTGTG-GTGTGTGCGCGCGGTGGCGGGGTGTGTGTGCTGTGCTG 921
QY 335 CTGGCGGCGGCGCTTCCAGCGGTGCCAGCTGCTCAGCGTGTGTGTACTGCGCGGCG 394
Db 920 TGGCGGCGCGCGCTGTGGCGCGCGCGAGCGCGACTCCCGGCGCGCGACGCGCGC 861
QY 395 GCGCAGCGCGGCGGACACAGCAAGCTTCCTGCTGCGCGACCCCTGTGATGACCTGAC 454
Db 860 GCGCGGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 801
QY 455 ACCCGGCAAGCGCTGTCTGAGTGTCTGGGCGCTGTCCAGGAGGACACACGCGCGACTTG 514
Db 800 CGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 741
QY 515 GCGGAGTTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 574
Db 740 GCGGCGGAGGCGGCGCTGTGCGG-CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 682
QY 575 GCGAGCGGCTGAGTGGGCTGCGCACAGTGTGCGCG 614
Db 681 CCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 642

RESULT 8
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 2.7%; Score 70.4; DB 3; Length 4411529;
Best Local Similarity 47.7%; Pred. No. 3.5e-05;
Matches 267; Conservative 0; Mismatches 291; Indels 2; Gaps 2;

QY 91 CGGGTCTCGCTGGGCGCATGCTCCGCGTGGCGCTTGTGCTGAGCTTCCCGACTG 150
Db 3944636 CGGCCCCACCGTGGCGCGCTTGCATTGCTCTCCGCGCGCGTTCGCGC 3944577
QY 151 CACCTTGGCTCACTTCGCGCTAGGCGCGCGAGCGCGCGCGAGCGAGCGCGCGCGCG 210
Db 3944576 CATTGCGGCGGACCGCGCGCTGCGCGCGTGGCGCGCGTTCGCGCGGATCCGCGACCC 3944517
QY 211 CGCGTGGCGGCGCTTGGGGCCCCCGGAGCGCAGTACTGCTGTGCGTGCCTGTGAC 270
Db 3944516 GGGGTGCGCGAGCGCGCGCGTGCCTGCGCGTGCCTGCGCGCGCGCGCGCGCGTCCC 3944457
QY 271 CCGGAGCGCGGTGCGGGCGCGGGTTCGCGGCGCGCGCGGTGCACACAGCGCTGTGCA 330
Db 3944456 CGCGCGGTGCGCGCGGTGCGCGCATCGCGCAATT-GCCGCGCGCAGCGCGGTGCGG 3944398
QY 331 CCAGTGTGCGCGCGCGCTTCAGAGGTGCCAGTGTCTCAGGCTGCTGCTACTGTGCC 390
```

```
Db 3944397 CCGGCGCGCGCGACTTGCCTGCGCGCGGTACCGCGCGTTCGCGCGCGTTCGCGCCA 3944338
QY 391 GGGCGCGCAGCGCGCGCGCGCACAGCAGCTTCTGTGCGCGGACCCCTGTGATGACCC 450
Db 3944337 TCGCGCGCAGTGC CGCGCGCAGGCTGCGATGCCGCGCGCGCGTGC-GCCGCTGCC 3944279
QY 451 TGACACCGCGCAAGCGCTGCTCAGCTGTGTGGCGCTGCCAGAGGACACAGCGCGCGCA 510
Db 3944278 GCGGTTCGCGCGGTGCCCGCTTGC CGCGCGGTGCTCAGTGTGAGTGGCGCGTTCGCG 3944219
QY 511 CTTGGCGAGTTCGAGGCGCGACCCGCGCGCGCTGTGGCAGCGCTCTGGAGGTGCA 570
Db 3944218 GATGGCGCGTTCGCGCGCATCGCGCGCGCGCGCTTCCCGCGCGTGCCTGTGCGCG 3944159
QY 571 AGACGCGAGCGCGCTGCGAGTGGGTGCGCACAGTGTGTCGCGCTCCCGAGCGCGCGCT 630
Db 3944158 ACGGTCAAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTCGCGCGGT 3944099
QY 631 GCACCGGTGTGCGCACT 650
Db 3944098 GCCACCGTCACGCGCGCGCT 3944079

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.7%; Score 68.8; DB 3; Length 4403765;
Best Local Similarity 47.5%; Pred. No. 8.5e-05;
Matches 266; Conservative 0; Mismatches 292; Indels 2; Gaps 2;

QY 91 CGGGTCTCGCTGGGCGCATGCTCCGCGTGGCGCTTGTGCTGAGCTTCCCGACTG 150
Db 3938183 CGGCCCCACCGTGGCGCGCTTGCATTGCTCTCCGCGCGCGTTCGCGCGTTCGCGC 3938124
QY 151 CACCTTGGCTCACTTCGCGCTAGGCGCGCGACCGCGCGCGAGCGAGCGCGCGCGCG 210
Db 3938123 CATTGCGGCGCGCACCGCGCGCTGCGCGCGTTCGCGCGCGGATCCGCGACCC 3938064
QY 211 CGCGTGGCGGCGTGTGGGGCCCCCGGAGCGCAGTACTGCTGTGCGTGGTGCCTGTGAC 270
Db 3938063 GGGGTGCGCGAGGCGCGCGTTCGCGCGGTTCGCGCGCGTTCGCGCGCGCGCGCGTCC 3938004
QY 271 CCGGAGCGCGCTGCGGGCGCGGTTCGCGCGCGCGGTTCGCGCGCGCTGTGAC 330
Db 3938003 CGCGCGGTGCGCGCGGTTCGCGCGCATCGCGCGCAATT-GCCGCGCGCAGCGCGGTGCGG 3937945
QY 331 CCAGTGTGCGCGCGCGCTTCAGCGGTGCCAGTGTCTCAGGCTGCTGCTACTGTGCC 390
Db 3937944 CCGGCGCGCGCGACCTTCGCGCGCTTCGCGCGGTTCGCGCGGTTCGCGCGTTCGCGCGCA 3937885
```



```
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic"
US-08-642-255-50

Query Match      2.6%; Score 67; DB 1; Length 756;
Best Local Similarity 45.3%; Pred. No. 2.4e-06;
Matches 287; Conservative 0; Mismatches 340; Indels 6; Gaps 1;

QY 76 GCTGCTCGGGGCGGGGCTGCGCTGGGCGCATGGCTCGCGCTGCGCTTGGTCTCT 135
DB 1 GGTGCTCAGCGCCAGCAGAGTTCGAAGGGCGCGCATGGCCAGCAGGCCCGAAAGGTGCG 60
QY 136 GGAGCTTCCCGACTGCACCTTGGCTCAGTTCGCTTGGCTAGGCGCGGAGCGCCCGGGCAGCG 195
DB 61 CTGGAGACCGGTGGTCCACCGGTGCTCCGGGACCTGCAGGCGCGCCAGGTGCGCTGGA 120
QY 196 AAGCGCCCGACCGCGCTGGCGCGCTGTGGGCGCCCGGAGCGCGAGCTACTCGCT 255
DB 121 CCGGTGCTGTCACCGGGTGTCTCGGACCTGCAGGCGCCCGCAGGTGCGCTGAGCGGCT 180
QY 256 GTGCGTGGCGGCGGCGGCGGCGGCTGCGGGGCGCGGCTGCGGGCGGCGGCTGCA 315
DB 181 GGTCCACCGGTGCTCCGGGACTGCAGGCGCGCGAGGTGCGCTGGAACGGGTGCTCA 240
QY 316 CCAGCGCTGCTGCACCGAGTGGCGCGGCGGCGGCGGCTTCCAGCGGTGCGGCTCAGGCT 375
DB 241 CCGGTGCTCCGGGACTGCAGGCGCGCGAGGTGGCGCTTGGACCGGCTGGTCCACCGGT 300
QY 376 GCTCTGCTACTGCCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 435
DB 301 GCTCCGGGACCTGCAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 436 CCGCTTGGATGACCTGCACCGCGCAAGCGTGTGCTGAGCTGCTGGGCGGCTGCCAGGA 495
DB 361 GACCTTGACGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 496 GGCACACCGCGCGCACTTGGGCGAGTTCAGAGCGCGACCGCGCGGCGGCGGCGGCGG 555
DB 421 GCAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 556 CTTCTGGGAGGTGCAAGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 613
DB 481 CCGCGAGGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 614 ----GTCGCGGAGCGCGCGCTGCACCGGTGGTGCGAGACTTGCACAGTTGCCAGTTGCTT 669
DB 541 GGTGCGCTTGGACCGGCTGGTCCACCGGTGTCTCCGGGACCTGCAGGCGCGGCGGCGG 702
QY 670 CCGGACCGGGAAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 702
DB 601 CTGGACCGGCTGGTCCACCGGTGCTCCGGGA 633

RESULT 12
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
; FILE REFERENCE: 0867/0905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match      2.5%; Score 64; DB 4; Length 1926;
Best Local Similarity 46.3%; Pred. No. 2.1e-05;
Matches 211; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 25 GCGCGCCATGGGCTTTCGCGCGCGGCTCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 84
DB 719 GCGCGCTGCTGCTGCGCGGCTCTGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 85 GCGCGCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 144
DB 659 GCTGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 145 CGACTGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 204
DB 599 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 205 CGACCGCGCGCTGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 264
DB 539 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 265 CTTGACCGCGGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 324
DB 479 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 325 GCTGCACAGCTGCGCGCGGCGGCTTCCAGCGGTGCGGCGGCGGCGGCGGCGGCGGCGG 384
DB 419 CTTGCGGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 385 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 444
DB 359 CTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 445 TGACCTGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 299 GCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 264

RESULT 13
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; FILE REFERENCE: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein-Barr virus
; NAME/KEY: CDS
US-09-410-399-3

Query Match      2.5%; Score 64; DB 4; Length 1926;
Best Local Similarity 46.3%; Pred. No. 2.1e-05;
Matches 211; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 25 GCGCGCCATGGGCTTTCGCGCGCGGCTCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 84
```



```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match      2.5%; Score 64; DB 3; Length 2580;
Best Local Similarity 46.3%; Pred. No. 2.4e-05;
Matches 211; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 25 GGCGCCCATGGCTTCGCGCGCGGCTCCTCGCGGGGCACATGTCGGGGCCGCTGCTCGG B4
Db 1102 GCCCTCTCTGCTCTGCGCCCTCTGCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTCT 1043
QY 85 GGGCGCGGGGTCGTGCGTGGGCCATGGCTCGCGCGTTCGTCGCGGCTTCCTGAGCTTCC 144
Db 1042 GCTCTGCGCCCTCTGCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCTCTGCGCCCT 983
QY 145 CGACTGCACCCCTGGCTCACTTGGCCCTAGCGCGGAGCGCCCGCGGCGAGCGCCCGC 204
Db 982 CCTCTGCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCG 923
QY 205 CGACCCCGCGCTGGCGCGCTGTGGGGCCCGCGAGCGAGCTACTGCTGTGGGTGCC 264
Db 922 CCTCTCTGCTCTGCGCCCTCTCTGCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTCT 863
QY 265 CGTGACCCCGAGCGCGGCTGCGGGGCCCGGGTCCGGGGCGCGCGCTGCACCGCGCT 324
Db 862 CCGTCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCGCCCT 803
QY 325 GCTGCACACAGCTGCGCGCGGCGCCCTTCCAGCGGTGCCAGCTGCTCAGCTGCTGCTA 384
Db 802 CTTGCCCTCTCTGCGCCCTCTCTGCTCTGCGCCCTCTCTGCTCTCTGCGCCCTCTCTGCC 743
QY 385 CTGCGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 444
Db 742 CCTCTGCGCCCTCTGCGCCCTCTCTGCTCTGCGCCCTCTCTGCGCCCTCTCTGCGCCCTCT 693
QY 445 TGACCTGTGACACCGGCGAGCGGCTGCTGAGCTGCT 480
Db 682 GCTCTGCGCCCTCTGCGCCCTCTGCTCTGCTGCTGTT 647
```

Search completed: October 22, 2004, 20:24:32
Job time : 235 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 07:31:36 ; Search time 1225 Seconds
(without alignments)
11017.356 Million cell updates/sec

Title: US-10-681-223-1
Perfect score: 2571
Sequence: 1 cggggccagggcggtcg.....aggaggggttttgcacc 2571

Scoring table: IDENTITY NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2571	100.0	2571	10	Adc63369 Human cDN
2	2571	100.0	2571	10	Abx93989 cDNA enco
3	2571	100.0	2571	12	Adm49355 Human kin
4	2380.6	92.6	3060	12	Adp84474 Human bre
5	2288	89.0	2307	12	Adi40904 Human kin
6	2276	88.5	2856	9	Ada38100 Thymidine
7	2276	88.5	2856	9	Adc39135 Novel hum
8	2276	88.5	2856	12	Adh42558 Novel hum
9	2274.4	88.5	2856	12	Adm10601 Human thy
10	2125.2	82.7	2645	8	Abx63720 Human cDN
11	1704.8	66.3	2331	10	Adc39137 Novel hum
12	1704.8	66.3	2331	12	Adh42564 Novel hum
13	1694.2	65.9	2566	10	Adf76970 Novel hum
14	1579.6	61.4	2199	5	Abv26811 Human pro
15	1579.6	61.4	2199	5	Abv20963 Human pro
16	1579.6	61.4	2199	5	Abv22074 Human pro
17	1579.6	61.4	2199	5	Abv27911 Human pro
18	1459.6	56.8	1535	9	Aad58542 Human thy
19	1193.8	46.4	1720	12	Adn05290 Antipsori
20	1193.8	46.4	1720	12	Ado24371 Human DNA
21	1130.4	44.0	20966	10	Adc63371 Human gen

22	1130.4	44.0	20966	10	ABX93990	Abx93990 Genomic D
23	1130.4	44.0	20966	12	ADNA9357	Adn49357 Human kin
24	1035.4	40.3	1579	6	ABL89542	Abi89542 Human pol
25	797.2	31.0	1188	12	ADJ75891	Adj75891 Marker ge
26	690.2	26.8	757	9	AAD58559	Aad58559 Human thy
27	674.8	26.2	759	9	AAD58558	Aad58558 Human thy
28	646.6	25.1	833	9	AAD58543	Aad58543 Human thy
29	641.8	24.2	844	9	AAD58550	Aad58550 Human thy
30	600.4	23.4	972	9	AAD58544	Aad58544 Human thy
31	569	22.1	728	9	AAD58552	Aad58552 Human thy
32	566.4	22.0	601	12	ADH42562	Adh42562 Novel hum
33	566.4	22.0	601	12	ADH42560	Adh42560 Novel hum
34	562	21.9	700	9	AAD58546	Aad58546 Human thy
35	562	21.9	700	9	AAD58553	Aad58553 Human thy
36	534	20.8	702	9	AAD58554	Aad58554 Human thy
37	533.2	20.7	795	9	AAD58547	Aad58547 Human thy
38	524.2	20.4	669	9	AAD58560	Aad58560 Human thy
39	506	19.7	1008	9	AAD58549	Aad58549 Human thy
40	493.8	19.2	928	9	AAD58557	Aad58557 Human thy
41	482.6	18.8	556	9	AAD58561	Aad58561 Human thy
42	480.4	18.7	674	9	AAD58548	Aad58548 Human thy
43	463.6	18.0	3292	6	ABH199757	Abi199757 Mouse lsc
44	456.6	17.8	633	9	AAD58545	Aad58545 Human thy
45	453.4	17.6	531	12	ADJ75005	Adj75005 Marker ge

ALIGNMENTS

RESULT 1
ADc63369
ID ADc63369 standard; cDNA; 2571 BP.
XX
AC ADc63369;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA encoding a novel kinase protein.
XX
KW Human; ss; gene; kinase; chromosome 2.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 32..11381
FT /*tag= a
FT /product= "Kinase"
XX
XX US2003087294-A1.
XX
XX 08-MAY-2003.
XX
XX 22-OCT-2002; 2002US-00277032.
XX
XX 31-OCT-2001; 2001US-00984880.
XX
XX (APPL-) APPLERA CORP.
XX
XX Wei M, Ketchum KA, Beasley EM, Difrancesco V;
XX
XX WPI: 2003-765435/72.
XX
XX P-PSDB; ADc63370.
XX
XX Novel isolated human kinase protein useful for drug screening assays, as
XX
XX a target for diagnosing disease, pharmacogenomic analysis, and for
XX
XX identifying compounds that modulate kinase activity.
XX
XX Claim 22; SEQ ID NO 1; 48pp; English.
XX
XX The invention relates to an isolated peptide consisting of or comprising
XX
XX an amino acid sequence chosen from an amino acid sequence of human kinase
XX
XX protein (ADc63370), an allelic variant of the kinase, orthologue or
XX
XX fragment. Also included are an isolated antibody that selectively binds

CC to the kinase, an isolated nucleic acid molecule encoding the kinase
CC proteins detailed above (appearing as the cDNA, AD63369, and gene,
CC ADC6371), a gene chip comprising the nucleic acids, a transgenic non-
CC human animal comprising the nucleic acids; a nucleic acid vector
CC comprising the nucleic acids, a host cell containing the vector,
CC detecting the presence of the kinase in a sample, detecting the presence
CC of the nucleic acids in a sample, a pharmaceutical composition comprising
CC a kinase binding agent, treating a disease or condition mediated by a
CC human kinase protein and identifying a modulator of the expression of the
CC kinase involves. The kinase is useful for identifying a modulator, for
CC identifying an agent that binds to the kinase, for biological assays
CC related to kinases, drug screening assays, identifying compounds that
CC modulate kinase activity, to screen a compound for the ability to
CC stimulate or inhibit interaction between the kinase protein and a
CC molecule that normally interacts with kinase protein, for competitive
CC binding assays to discover compounds that interact with kinase,
CC pharmacogenomic analysis and treating a disorder characterised by an
CC absence of inappropriate or unwanted expression of the protein, useful as
CC target spot diagnosing active protein activity, disease or predisposition
CC to disease. The nucleic acid is useful for assessing expression in
CC disease states, diagnostic tools as an immunological marker for aberrant
CC protein for inhibiting protein function, tissue typing. The nucleic acids
CC are useful for facilitating protein trafficking or facilitate
CC manipulation of a protein for assay or production, probes, primers,
CC chemical intermediates, constructing recombinant vectors, expressing
CC antigenic portions of the proteins, for determining chromosomal positions
CC of the nucleic acid molecules, for making vectors containing the gene
CC regulatory regions, as hybridisation probes for determining the presence,
CC level, form and distribution of nucleic acid expression. The gene
CC encoding the kinase is located on human chromosome 2. The present
CC sequence is the cDNA encoding the kinase.

XX
SQ Sequence 2571 BP; 561 A; 711 C; 675 G; 624 T; 0 U; 0 Other;

Query Match 100.0%; Score 2571; DB 10; Length 2571;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCCGAGGCGCGTCTGAGCGGCCATGCGCTTCGCGCGCGCTCTGCGCGG 60
DB |||||
QY 1 CGGGCCGAGGCGCGTCTGAGCGGCCATGCGCTTCGCGCGCGCTCTGCGCGG 60
DB |||||
QY 61 GCCACTGTCTGCGCGCGCTCTGCGCGCGCGCGGCTCTGCGCGCGCGCTCTGCGCGG 120
DB |||||
QY 61 GCCACTGTCTGCGCGCGCTCTGCGCGCGCGCGGCTCTGCGCGCGCGCTCTGCGCGG 120
DB |||||
QY 121 GTGCGCTCTGCTGAGCTTCCGACTGCGACCTGCGCTACTGCGCGCGCGCTCTGCGCGG 180
DB |||||
QY 121 GTGCGCTCTGCTGAGCTTCCGACTGCGACCTGCGCTACTTCCGCGCGCGCTCTGCGCGG 180
DB |||||
QY 181 CSCCCCGGCGAGCGAGCG 240
DB |||||
QY 181 CGCGCCCGGCGAGCGAGCG 240
DB |||||
QY 241 GCGAGCTACTGCTGTGCGTCTGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB |||||
QY 241 GCGAGCTACTGCTGTGCGTCTGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB |||||
QY 301 GCGCGCGCGCTGCGACGAGCGCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB |||||
QY 301 GCGCGCGCGCTGCGACGAGCGCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB |||||
QY 361 CCAGCTGTCTGAGCTCTGCTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB |||||
QY 361 CCAGCTGTCTGAGCTCTGCTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB |||||
QY 421 CTTCTGTCTGCG 480
DB |||||
QY 421 CTTCTGTCTGCG 480
DB |||||
QY 481 GGGCGCTCTGCGAGGCG 540
DB |||||
QY 481 GGGCGCTCTGCGAGGCG 540
DB |||||

QY 541 CCAGCTGTGCGAGCGCGCTCTGCGAGAGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 600
DB |||||
QY 541 CCAGCTGTGCGAGCGCGCTCTGCGAGAGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 600
DB |||||
QY 601 ACAGTGTGCGCGCGCTCTGCGAGAGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 660
DB |||||
QY 601 ACAGTGTGCGCGCGCTCTGCGAGAGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 660
DB |||||
QY 661 CGTGTGTCTTCCCGGAGCG 720
DB |||||
QY 661 CGTGTGTCTTCCCGGAGCG 720
DB |||||
QY 721 TCTGTAGCGCGCGCGAGTGTCTGAGTGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 780
DB |||||
QY 721 TCTGTAGCGCGCGCGAGTGTCTGAGTGTGCAAGCGCAGCGCGCTCTGAGTGGGCTGCGG 780
DB |||||
QY 781 AAAGTTCAGGTTGTGTCATCGAAGGACTGATGCGCGGCTGCAAGCGCGCTCTGAGTGGGCTGCGG 840
DB |||||
QY 781 AAAGTTCAGGTTGTGTCATCGAAGGACTGATGCGCGGCTGCAAGCGCGCTCTGAGTGGGCTGCGG 840
DB |||||
QY 841 GTCAGTGTGCGAGTTCACCTTAAGGCTGTCTCTTAAAGTCAACCGCTCTTGTGATGGCGA 900
DB |||||
QY 841 GTCAGTGTGCGAGTTCACCTTAAGGCTGTCTCTTAAAGTCAACCGCTCTTGTGATGGCGA 900
DB |||||
QY 901 GTGAGGAGAGTCTTGTGATGAGCAACACTATCATTAGAGAGCTTTTACTCTTTGGG 960
DB |||||
QY 901 GTGAGGAGAGTCTTGTGATGAGCAACACTATCATTAGAGAGCTTTTACTCTTTGGG 960
DB |||||
QY 961 CAATTATATTTGTGGCTCCGAAATAGCTAAAGAACTGCAAAATCTCTCTGTGATCTAGA 1020
DB |||||
QY 961 CAATTATATTTGTGGCTCCGAAATAGCTAAAGAACTGCAAAATCTCTCTGTGATCTAGA 1020
DB |||||
QY 1021 CAGGTACTGGCAGCAGCGCGCACCTATGCCATAGCCACTGAGGTGAGTGGGCGTCTCCA 1080
DB |||||
QY 1021 CAGGTACTGGCAGCAGCGCGCACCTATGCCATAGCCACTGAGGTGAGTGGGCGTCTCCA 1080
DB |||||
QY 1081 GCACCTGCG 1140
DB |||||
QY 1081 GCACCTGCG 1140
DB |||||
QY 1141 CCTTATCTCTGCTGCTCAGCTGTGAGTCTCTGAGGAGAGGTTCAGAGGCTGTCAGGCGCGG 1200
DB |||||
QY 1141 CCTTATCTCTGCTGCTCAGCTGTGAGTCTCTGAGGAGAGGTTCAGAGGCTGTCAGGCGCGG 1200
DB |||||
QY 1201 CATGAGAGAGCAGGAG 1260
DB |||||
QY 1201 CATGAGAGAGCAGGAG 1260
DB |||||
QY 1261 AGAATATGCTTACCGAGCGAGTGGAGAACTCTGCGTCCCATGTTGATGCGAGCGCGCTC 1320
DB |||||
QY 1261 AGAATATGCTTACCGAGCGAGTGGAGAACTCTGCGTCCCATGTTGATGCGAGCGCGCTC 1320
DB |||||
QY 1321 CAGAGAAAGGTCTCTGAGCAGAGTATTAAGCCTAATCCAGAAATAGTCTTGTAGTGAACCGTA 1380
DB |||||
QY 1321 CAGAGAAAGGTCTCTGAGCAGAGTATTAAGCCTAATCCAGAAATAGTCTTGTAGTGAACCGTA 1380
DB |||||
QY 1381 GTTACTCTGCGCGAGTGGCGACGCTCTAACTAGATTAGTGTGTTGAAACATCTCATCC 1440
DB |||||
QY 1381 GTTACTCTGCGCGAGTGGCGACGCTCTAACTAGATTAGTGTGTTGAAACATCTCATCC 1440
DB |||||
QY 1441 ACCATTTGTATGCAAGTGTCTCCAAATTTCTCTTACAGCATCTGTTGTTGCGCAGAA 1500
DB |||||
QY 1441 ACCATTTGTATGCAAGTGTCTCCAAATTTCTCTTACAGCATCTGTTGTTGCGCAGAA 1500
DB |||||
QY 1501 CTGAGAGACGAGCATCTTAAATTTTACTTACAGCATCTGTTGTTGCGCAGAA 1560
DB |||||
QY 1501 CTGAGAGACGAGCATCTTAAATTTTACTTACAGCATCTGTTGTTGCGCAGAA 1560
DB |||||
QY 1561 CGTCTATCAAGAGTCTCTCTCATCATGTTCCAGTGTGAGGCGCGAGTGTCTTCTTCC 1620
DB |||||
QY 1561 CGTCTATCAAGAGTCTCTCTCATCATGTTCCAGTGTGAGGCGCGAGTGTCTTCTTCC 1620
DB |||||

QY 1621 TGGCATAGTAAACATTTCTTTGGAAACATATGTTTCACTTAATCACTACCAATATCTGGA 1680
DB 1621 TGGCATAGTAAACATTTCTTTGGAAACATATGTTTCACTTAATCACTACCAATATCTGGA 1680
QY 1681 AGACTGTCTTACTCAGACAGCACCCAGGTGTACAGACGACGACAGATCTTCCAGAT 1740
DB 1681 AGACTGTCTTACTCAGACAGCACCCAGGTGTACAGACGACGACAGATCTTCCAGAT 1740
QY 1741 CAGCAGGAGAGACCCCGGAGCCTCTGCTTCTCTACACTGGCATGCTGATGATCGTGAC 1800
DB 1741 CAGCAGGAGAGACCCCGGAGCCTCTGCTTCTCTACACTGGCATGCTGATGATCGTGAC 1800
QY 1801 ATGCCCAATTGGCTTCTTCCACATCTGGTTGCACTCTGATGATGGCTGCTGCATCT 1860
DB 1801 ATGCCCAATTGGCTTCTTCCACATCTGGTTGCACTCTGATGATGGCTGCTGCATCT 1860
QY 1861 CCTCAGTCCCAAAATCTAGTAGCCAAAGTTCTCTGCAAGAGCTCTCTATGTCTCTGGC 1920
DB 1861 CCTCAGTCCCAAAATCTAGTAGCCAAAGTTCTCTGCAAGAGCTCTCTATGTCTCTGGC 1920
QY 1921 TGCCCAAGGGACACCTCTCGAGAGCCATTTTGGGTAAAGGAACACTTACAAAGAAAGGCAT 1980
DB 1921 TGCCCAAGGGACACCTCTCGAGAGCCATTTTGGGTAAAGGAACACTTACAAAGAAAGGCAT 1980
QY 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTGATAGGCTTCTGATGCTCATTCATTAAGAC 2040
DB 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTGATAGGCTTCTGATGCTCATTCATTAAGAC 2040
QY 2041 ATTCAGCAAGATGCTCCAACTGCAATATATACCAACCTTCTCTGAATATATATTTGCTT 2100
DB 2041 ATTCAGCAAGATGCTCCAACTGCAATATATACCAACCTTCTCTGAATATATATTTGCTT 2100
QY 2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGCAATTTTCCA 2160
DB 2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGCAATTTTCCA 2160
QY 2161 TCTGAATCTGGATGATATCATTTAGCCAAATCAGTAAATTTATTTATTTATTAATCTATACAT 2220
DB 2161 TCTGAATCTGGATGATATCATTTAGCCAAATCAGTAAATTTATTTATTTATTAATCTATACAT 2220
QY 2221 AATATGTTTCTCAGCATAGGAGCTATGATTTCAATTAATTAAGTGGAGTCAAAAGCCTA 2280
DB 2221 AATATGTTTCTCAGCATAGGAGCTATGATTTCAATTAATTAAGTGGAGTCAAAAGCCTA 2280
QY 2281 AATGCAATGTTTGTGTGATTTTTCATTAACAAACTTAAATTTGCTTGTAAATAGTT 2340
DB 2281 AATGCAATGTTTGTGTGATTTTTCATTAACAAACTTAAATTTGCTTGTAAATAGTT 2340
QY 2341 CAACTGGATCTTGGAGTGGGATTTCTTGTAATATCTTGCACTGAATGCTCATGAT 2400
DB 2341 CAACTGGATCTTGGAGTGGGATTTCTTGTAATATCTTGCACTGAATGCTCATGAT 2400
QY 2401 TACATATGAATCGCTTTTACATATCTTTAGACAGAAAAAGTAGTGTAGTGGGGGAA 2460
DB 2401 TACATATGAATCGCTTTTACATATCTTTAGACAGAAAAAGTAGTGTAGTGGGGGAA 2460
QY 2461 ATTATAGACTTGTGTGATTTTATAGGAGTAGCTGTCTTTATACATACTCAAGCCCTG 2520
DB 2461 ATTATAGACTTGTGTGATTTTATAGGAGTAGCTGTCTTTATACATACTCAAGCCCTG 2520
QY 2521 AAGCCTTGATGTCCTGCGAGCTCGCACTAAAGAGGGGGCTTTTGACACC 2571
DB 2521 AAGCCTTGATGTCCTGCGAGCTCGCACTAAAGAGGGGGCTTTTGACACC 2571

RESULT 2
ABX93989
ID ABX93989 standard; cDNA; 2571 BP.
XX
AC ABX93989;
XX
DT 10-JUN-2003 (first entry)
XX

DE cDNA encoding novel human thymidylate kinase subfamily kinase.
KW Human; thymidylate kinase subfamily; kinase; therapeutic agent;
KW immune response; kinase associated disorder; tissue typing;
KW pharmacogenomic analysis; gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
5'UTR 1..31
FT /*tag= a
FT CDS 32..1381
FT /*tag= b
FT /product= "Kinase"
FT 1382
FT /*tag= c
XX US6489153-B1.
PN 03-DEC-2002.
XX 31-OCT-2001; 2001US-00984880.
XX 31-OCT-2001; 2001US-00984880.
XX (PEKE) PE CORP NY.
XX Wei M, Ketchum KA, Beasley EM, Difrancesco V;
XX WPI; 2003-327246/31.
XX P-PSDB; ABU08678.
XX New human kinase polypeptides and nucleic acid molecules, useful in the
PT development of human therapeutic targets, or for diagnosing or treating a
PT disorder associated with the aberrant expression of the protein or
PT nucleic acid molecule.
XX Claim 5; Fig 1A-C; 49pp; English.
XX The invention describes a new isolated nucleic acid molecule encoding a
CC kinase of the thymidylate kinase subfamily. The human kinase polypeptides
CC and nucleic acid molecules are useful as models for the development of
CC human therapeutic targets, aid in the identification of therapeutic
CC proteins, or serve as targets for the development of human therapeutic
CC agents that modulate kinase activity in cells and tissues that express
CC the kinase. The kinase proteins are also useful in raising antibodies or
CC eliciting another immune response, as reagents in assays designed to
CC quantitatively determine levels of the protein in biological fluids, as
CC markers for tissues in which the corresponding protein is preferentially
CC expressed, or in screening a compound for the ability to stimulate or
CC inhibit interaction between the kinase protein and a molecule that
CC normally interacts with the kinase protein. The kinase proteins and
CC nucleic acid molecules can also be used in providing a target for
CC diagnosing a disease or a predisposition to a disease mediated by the
CC polypeptide or nucleic acid molecule, or for treating a disorder
CC associated with the absence of, inappropriate, or unwanted expression of
CC the protein or nucleic acid molecule. The antibodies are useful for
CC tissue typing, in isolating or detecting kinase proteins, or in
CC pharmacogenomic analysis. This sequence encodes the novel human
CC thymidylate kinase subfamily kinase
SQ Sequence 2571 BP; 561 A; 711 C; 675 G; 624 T; 0 U; 0 Other;
Query Match 100.0%; Score 2571; DB 10; Length 2571;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGGCGGAGGCGGCGCTGCTGAGCGGCCCATGCGCTCGCCCGCGCTCTCGCGG 60
DB 1 CGGGGCGGAGGCGGCGCTGCTGAGCGGCCCATGCGCTCGCCCGCGCTCTCGCGG 60
QY 61 GCCACTGTGCGGGCGCTGCTCGGGCGGCGGGGTCTCGCTGGGGCCATGCTCCGCC 120

Db 61 GCCACTGTGCGGGCCGCTGCTCGGGCGGCGGGGTCTCGGCTGGGGCCATGGCTCCGCC 120
QY 121 GTGCCGCTTCGTCTGGAGCTTCCGACTGCACTTGGGTCACTTCGCCCTTAGGGGCCGA 180
Db 121 GTGCCGCTTCGTCTGGAGCTTCCGACTGCACTTGGGTCACTTCGCCCTTAGGGGCCGA 180
QY 181 CGCCCCCGGCGAGCGCCCCCGAACCCCGCTGCGGGCGCTGTCTGGGGCCCCCGGA 240
Db 181 CGCCCCCGGCGAGCGCCCCCGAACCCCGCTGCGGGCGCTGTCTGGGGCCCCCGGA 240
QY 241 GCGCACTACTGGCTGTGGTGGCTGACCCCGGACCGCGGTGCGGGGCCCGGGTCCG 300
Db 241 GCGCACTACTGGCTGTGGTGGCTGACCCCGGACCGCGGTGCGGGGCCCGGGTCCG 300
QY 301 GCGGCGCGGCTGCACACAGCGCTGTGCACACAGCTGCGCGCGGCCCTTCCAGCGGTG 360
Db 301 GCGGCGCGGCTGCACACAGCGCTGTGCACACAGCTGCGCGCGGCCCTTCCAGCGGTG 360
QY 361 CCAGCTGCTCAGGCTGCTGTGCTACTGCCGCGGCGCCAGGCGGGCGGCGACAGCAAG 420
Db 361 CCAGCTGCTCAGGCTGCTGTGCTACTGCCGCGGCGCCAGGCGGGCGGCGACAGCAAG 420
QY 421 CTTCTGCTGCGCGACCCCGCTGATGACCTGACACCCCGGCAAGCGCTGCTCAGCTGCT 480
Db 421 CTTCTGCTGCGCGACCCCGCTGATGACCTGACACCCCGGCAAGCGCTGCTCAGCTGCT 480
QY 481 GGGCGCTCCGAGGAGGACACGCGCCGCACTTTGGCGGAGTTGAGGCCGACCCCGCGG 540
Db 481 GGGCGCTCCGAGGAGGACACGCGCCGCACTTTGGCGGAGTTGAGGCCGACCCCGCGG 540
QY 541 CCAGCTGTGCGAGCGCTCTGGGAGGTGCAAGACGCGAGCGGCTGCAAGTGGGCTGGCG 600
Db 541 CCAGCTGTGCGAGCGCTCTGGGAGGTGCAAGACGCGAGCGGCTGCAAGTGGGCTGGCG 600
QY 601 ACAGTGTGCGCGTCCCGAGCGCCCGCTGCAACCGGTTGGTCCAGACTTGGCCAGTTC 660
Db 601 ACAGTGTGCGCGTCCCGAGCGCCCGCTGCAACCGGTTGGTCCAGACTTGGCCAGTTC 660
QY 661 CGTGTCTTCCGAGCGGGAAGCGCCGCGGCGGTTTGGAGGAGTGTACCTCTTTAT 720
Db 661 CGTGTCTTCCGAGCGGGAAGCGCCGCGGCGGTTTGGAGGAGTGTACCTCTTTAT 720
QY 721 TCTGAAGCCCGGAGTGTCTTGAACCTGTGACACAGTGCACCAAGTGCACCAAGATCCAGAAAG 780
Db 721 TCTGAAGCCCGGAGTGTCTTGAACCTGTGACACAGTGCACCAAGTGCACCAAGATCCAGAAAG 780
QY 781 AAAGTTCAGGTTGTGCCATCGAAGGACTGGATGCCCGGTTAAACACACGCGTACCCA 840
Db 781 AAAGTTCAGGTTGTGCCATCGAAGGACTGGATGCCCGGTTAAACACACGCGTACCCA 840
QY 841 GTCAGTGGCAGATTCACTTAAGCTGTCTTAAAGTCAACACCTCTTGCATTGGCCA 900
Db 841 GTCAGTGGCAGATTCACTTAAGCTGTCTTAAAGTCAACACCTCTTGCATTGGCCA 900
QY 901 GTGGAGGAAGTCTTTGATGATGAACCAACTATCATTTAGAAGGCTTTTACTCTTTGG 960
Db 901 GTGGAGGAAGTCTTTGATGATGAACCAACTATCATTTAGAAGGCTTTTACTCTTTGG 960
QY 961 CAATTATATGTGGCTCCGAAATAGCTAAAGAACTGCAAACTCTCTGTGATTGTAGA 1020
Db 961 CAATTATATGTGGCTCCGAAATAGCTAAAGAACTGCAAACTCTCTGTGATTGTAGA 1020
QY 1021 CAGGTACTGGCAGACGCGCACCTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCA 1080
Db 1021 CAGGTACTGGCAGACGCGCACCTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCA 1080
QY 1081 GCACCTGCCCGGAGCCCATCACCCTGTGTACCACTGACAGTGGCGAGGACCTGCTCAAACTGA 1140
Db 1081 GCACCTGCCCGGAGCCCATCACCCTGTGTACCACTGACAGTGGCGAGGACCTGCTCAAACTGA 1140
QY 1141 CCTTATCTGCTGCTCACTGTGAGTCTGTGAGGAGAGGTTGCGAGGCTGCGAGGCGGG 1200
Db 1141 CCTTATCTGCTGCTCACTGTGAGTCTGTGAGGAGAGGTTGCGAGGCTGCGAGGCGGG 1200

QY 1201 CATGAGAAGACACAGGAAGAAGCAGAACTTGAGGCCAACAGTGTGTTTCTGTCAAAGGT 1260
Db 1201 CATGAGAAGACACAGGAAGAAGCAGAACTTGAGGCCAACAGTGTGTTTCTGTCAAAGGT 1260
QY 1261 AGAAATGTCTTACCAGCGATGGAAATCCTGGCTGCCATGTGGTTGATGCCAGGCCCTTC 1320
Db 1261 AGAAATGTCTTACCAGCGATGGAAATCCTGGCTGCCATGTGGTTGATGCCAGGCCCTTC 1320
QY 1321 CAGAGAAAAGTGCTTCGACAGTAGTATTAAGCCTAATCCAGAATAGTTTTAGTGAAACCGTA 1380
Db 1321 CAGAGAAAAGTGCTTCGACAGTAGTATTAAGCCTAATCCAGAATAGTTTTAGTGAAACCGTA 1380
QY 1381 GTTACTCTGGCAGGTGCCACGCTTAACTAGATTAGATCTGTTCAAACTACTACATCC 1440
Db 1381 GTTACTCTGGCAGGTGCCACGCTTAACTAGATTAGATCTGTTCAAACTACTACATCC 1440
QY 1441 ACCATTGTGTATGCAAGTGTCCCAAATTTCTGTCTACAGCATGTGTGTGGCAGAAAA 1500
Db 1441 ACCATTGTGTATGCAAGTGTCCCAAATTTCTGTCTACAGCATGTGTGTGGCAGAAAA 1500
QY 1501 CTGGAGACAGGCACTTAAATTTTACTTCAGCCATCGTACCTCTTCTGACTGATGGACC 1560
Db 1501 CTGGAGACAGGCACTTAAATTTTACTTCAGCCATCGTACCTCTTCTGACTGATGGACC 1560
QY 1561 CGTCACTCAAAAGGTCCCTCTCATGTTCAGTGAGAGGCCAGGATTTGCTTTCTTCC 1620
Db 1561 CGTCACTCAAAAGGTCCCTCTCATGTTCAGTGAGAGGCCAGGATTTGCTTTCTTCC 1620
QY 1621 TGGCATAGTAAACATTTTCTTGGAACTATGTTTCACTTAATCACTACCAATATCTGGA 1680
Db 1621 TGGCATAGTAAACATTTTCTTGGAACTATGTTTCACTTAATCACTACCAATATCTGGA 1680
QY 1681 AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAGCAGCAGACAGATCTCCAGAT 1740
Db 1681 AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAGCAGCAGACAGATCTCCAGAT 1740
QY 1741 CAGCAGGAGACCCCGGAGCCTCTGCTTCTCTACACTGCGATGCTGATGAGATCTGAC 1800
Db 1741 CAGCAGGAGACCCCGGAGCCTCTGCTTCTCTACACTGCGATGCTGATGAGATCTGAC 1800
QY 1801 ATGCCACACTTGGCTCTTTCCACATCTGGTTGCACTCGTCACTGATGGGCTCGCTGATCT 1860
Db 1801 ATGCCACACTTGGCTCTTTCCACATCTGGTTGCACTCGTCACTGATGGGCTCGCTGATCT 1860
QY 1861 CCTCAGTCCCAAAATTTCTAGTAGCCAGGTTCCTGACAGGCTGTCTATGTCTCTGGC 1920
Db 1861 CCTCAGTCCCAAAATTTCTAGTAGCCAGGTTCCTGACAGGCTGTCTATGTCTCTGGC 1920
QY 1921 TGCCCAAGGACACTCTCGAGGCCATTTTGGGTAAGGAACTTACATAAGGAC 1980
Db 1921 TGCCCAAGGACACTCTCGAGGCCATTTTGGGTAAGGAACTTACATAAGGAC 1980
QY 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTTGTAGAGCTTCTGATGCTCATTTATAAGAC 2040
Db 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTTGTAGAGCTTCTGATGCTCATTTATAAGAC 2040
QY 2041 ATTCAAGCCAGATGCTCCAACTGMAATATACCAACCTTCTGAAATTAATTTTGGCTT 2100
Db 2041 ATTCAAGCCAGATGCTCCAACTGMAATATACCAACCTTCTGAAATTAATTTTGGCTT 2100
QY 2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGCAATTTTCCA 2160
Db 2101 ATTTATATTTCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAAATGCAATTTTCCA 2160
QY 2161 TCTGAACCTGATGATCATTTAGCCAAATCCAGTAATTTATTTATTTAATCTATACAT 2220
Db 2161 TCTGAACCTGATGATCATTTAGCCAAATCCAGTAATTTATTTAATTTAATCTATACAT 2220
QY 2221 AATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAAGCTA 2280
Db 2221 AATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAAGCTA 2280


```
QY 961 CAATTATATCTGGCTCCGAATAGCTAAAGATCTGCCAATCTCCTGTGATGTAGA 1020
Db 961 CAATTATATCTGGCTCCGAATAGCTAAAGATCTGCCAATCTCCTGTGATGTAGA 1020
QY 1021 CAGGTACTGGCAGACAGCGGCACCTATGCCATAGCACTGAGGTGAGTGGGGGTCTCCA 1080
Db 1021 CAGGTACTGGCAGACAGCGGCACCTATGCCATAGCACTGAGGTGAGTGGGGGTCTCCA 1080
QY 1081 GCACCTGCCCCAGCCCATCACCCCTGTGTACCACTGGCCAGAGGACCTCTCAAACTGA 1140
Db 1081 GCACCTGCCCCAGCCCATCACCCCTGTGTACCACTGGCCAGAGGACCTCTCAAACTGA 1140
QY 1141 CTTTATCTCTGCTCTCACTGTGAGTCTCTGAGGAGAGTTGCGAGAGGCTCGAGGGCCGGGG 1200
Db 1141 CTTTATCTCTGCTCTCACTGTGAGTCTCTGAGGAGAGTTGCGAGAGGCTCGAGGGCCGGGG 1200
QY 1201 CATGGAGAAGACAGAGGAAGAGCAGAACTTGAGGCCAACAGTGTCTTTTCGTCAAAAGGT 1260
Db 1201 CATGGAGAAGACAGAGGAAGAGCAGAACTTGAGGCCAACAGTGTCTTTTCGTCAAAAGGT 1260
QY 1261 AGAATGTCTTACGACGGATGGAGAACTCTGTGCTGCCATGTGGTGTGATGCCAGCCCCCTC 1320
Db 1261 AGAATGTCTTACGACGGATGGAGAACTCTGTGCTGCCATGTGGTGTGATGCCAGCCCCCTC 1320
QY 1321 CAGAGAAAAGGTCTCGACACAGTATTAAAGCCCTAATCCAGAAATAGTGTAGTGAACCGTA 1380
Db 1321 CAGAGAAAAGGTCTCGACACAGTATTAAAGCCCTAATCCAGAAATAGTGTAGTGAACCGTA 1380
QY 1381 GTTACTCTGGCCAGGTGCCAGCTCTAACTAGATTAGATGTGTGTTGAAAACATCTACATCC 1440
Db 1381 GTTACTCTGGCCAGGTGCCAGCTCTAACTAGATTAGATGTGTGTTGAAAACATCTACATCC 1440
QY 1441 ACCATTTGTTATGCAAGTGTCCCAATTTCTGTTCTACAGCATGTGTGTGGCGAGAAA 1500
Db 1441 ACCATTTGTTATGCAAGTGTCCCAATTTCTGTTCTACAGCATGTGTGTGGCGAGAAA 1500
QY 1501 CTGGAGACAGGCATCTTAATTTTACTTTCAGCCATCGTACCTCTCTGACTGATGGAC 1560
Db 1501 CTGGAGACAGGCATCTTAATTTTACTTTCAGCCATCGTACCTCTCTGACTGATGGAC 1560
QY 1561 CGTCTACAAAAGTCCCTCTCATCATGTTCAGTGGAGGCGCAGCGATTGCTTTCTCC 1620
Db 1561 CGTCTACAAAAGTCCCTCTCATCATGTTCAGTGGAGGCGCAGCGATTGCTTTCTCC 1620
QY 1621 TGGCATAGTAAACATTTCTTGGAACATATGTTTCACTTAATCACTACCAAAATATCTGA 1680
Db 1621 TGGCATAGTAAACATTTCTTGGAACATATGTTTCACTTAATCACTACCAAAATATCTGA 1680
QY 1681 AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAAAGCAGACAGATCTCCAGAT 1740
Db 1681 AGACCTGTCTTACTCAGACAGCACAGGTGTACAGAAAGCAGACAGATCTCCAGAT 1740
QY 1741 CAGCAGGAGACCCCGAGCCTCTGCTTCTCTTACACTGGCATGCTGATGAGATCGTAC 1800
Db 1741 CAGCAGGAGACCCCGAGCCTCTGCTTCTCTTACACTGGCATGCTGATGAGATCGTAC 1800
QY 1801 ATGCCACATTTGGCTTCTTCCACATCTGGTTGCACTCGTCATGATGGCTCGCTGATCT 1860
Db 1801 ATGCCACATTTGGCTTCTTCCACATCTGGTTGCACTCGTCATGATGGCTCGCTGATCT 1860
QY 1861 CCTCAGTCCCAATTTAGTAGCCCAAGTGTCTCTGAGAGGCTGTATGTCTCTGGC 1920
Db 1861 CCTCAGTCCCAATTTAGTAGCCCAAGTGTCTCTGAGAGGCTGTATGTCTCTGGC 1920
QY 1921 TGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAAGAACTTACAAAGAGGCAT 1980
Db 1921 TGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAAGAACTTACAAAGAGGCAT 1980
QY 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTCATAAGAC 2040
Db 1981 TGATCTTGTCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTCATAAGAC 2040
```

```
QY 2041 ATTCAAGCCAGATGCTCCAACTGCAAAATATACCAACCTTCTCTGAATATATTTTGCCT 2100
Db 2041 ATTCAAGCCAGATGCTCCAACTGCAAAATATACCAACCTTCTCTGAATATATTTTGCCT 2100
QY 2101 ATTTATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2160
Db 2101 ATTTATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2160
QY 2161 TCTGAACCTGGATGCATATCAITTTAGCCANTCCAGTAATTTATTTATTTAACTATACAT 2220
Db 2161 TCTGAACCTGGATGCATATCAITTTAGCCANTCCAGTAATTTATTTATTTAACTATACAT 2220
QY 2221 AATATGTTTCTCCAGCATAGGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAACGCTA 2280
Db 2221 AATATGTTTCTCCAGCATAGGAGCTATGATTCATTAATTTAAAGTGGAGTCAAAACGCTA 2280
QY 2281 AATGCAATGTTTGTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2340
Db 2281 AATGCAATGTTTGTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2340
QY 2341 CAAGTGGATCTTGGAGTGGATTTCTTGTGTAATTTCTTGCTTGAATGTCTCATGAT 2400
Db 2341 CAAGTGGATCTTGGAGTGGATTTCTTGTGTAATTTCTTGCTTGAATGTCTCATGAT 2400
QY 2401 TACATATGAAATCGCTTTTGCACATATCTTTAGACAGAAAAGTAGCTGAGTGGGGGAA 2460
Db 2401 TACATATGAAATCGCTTTTGCACATATCTTTAGACAGAAAAGTAGCTGAGTGGGGGAA 2460
QY 2461 ATTATAGACTTGTGTGACTTTAGGAGTAGCTGTCTCTTTATACATACTCAGCCCTG 2520
Db 2461 ATTATAGACTTGTGTGACTTTAGGAGTAGCTGTCTCTTTATACATACTCAGCCCTG 2520
QY 2521 AAGCCTTGCATGCTCTCGAGCTCGCACTAAAGAGGGGGCTTTTGCACC 2571
Db 2521 AAGCCTTGCATGCTCTCGAGCTCGCACTAAAGAGGGGGCTTTTGCACC 2571
```

RESULT 4

ID ADP84474 standard; DNA; 3060 BP.

XX ADP84474;

XX AC

XX 09-SEP-2004 (first entry)

XX Human breast-specific protein coding sequence #73.

XX human; breast-specific protein; breast cancer; gene; ds.

XX Homo sapiens.

XX WO2004053077-A2.

XX 24-JUN-2004.

XX 05-DEC-2003; 2003WO-US038815.

XX 05-DEC-2002; 2002US-0431123P.

XX (DIAD-) DIADEXUS INC.

XX Macina RA, Turner LR, Sun Y, Chen H, Rodriguez M;

XX WPI; 2004-468848/44.

XX P-PSDB; ADP84606.

XX New breast specific nucleic acid molecules and polypeptides useful for
PT diagnosing, preventing or treating breast cancer, for producing
PT transgenic animals or cells, or for research purposes.

XX Claim 1; SEQ ID NO 73; 521pp; English.

XX The invention comprises the amino acid and coding sequences of human

CC breast-specific proteins. The DNA and protein sequences of the invention
CC are useful for the diagnosis, treatment and prevention of breast cancer.
CC The present DNA sequence encodes a human breast-specific protein of the
CC invention.

XX
SQ Sequence 3060 BP; 731 A; 758 C; 757 G; 814 T; 0 U; 0 Other;

Query Match 92.6%; Score 2380.6; DB 12; Length 3060;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2463; Conservative 0; Mismatches 9; Indels 7; Gaps 7;

QY	15	GGTGGCTGAGGGCCCATGGCTTCGCGCGCGGCTCTGCGCGGCGGCTGCGGGG	74
DB	1	GGTGGCTGAGGGCCCATGGCTTCGCGCGCGGCTCTGCGCGGCGGCTGCGGGG	60
QY	75	CGCTGCTGGGGCGGCGGGGCTGCGCTGGGGCCATGGCTTCGCGCGGCTGCGTC	134
DB	61	CGCTGCTGGGGCGGCGGGGCTGCGCTGGGGCCATGGCTTCGCGCGGCTGCGTC	120
QY	135	TGAGCTTCGCGACTGACCCCTGGCTCACTTCGCGCTAGGGCGGAGCCCGGGGAG	194
DB	121	TGAGCTTCGCGACTGACCCCTGGCTCACTTCGCGCTAGGGCGGAGCCCGGGGAG	180
QY	195	CAGAGCCCGGACCCCGGCTGGCGGCGCTGTTGGGGCCCGGAGCGAGCTACTCG	254
DB	181	CAGAGCCCGGACCCCGGCTGGCGGCGCTGTTGGGGCCCGGAGCGAGCTACTCG	240
QY	255	TGTGGCTGCGCGTGACCCCGGAGCGCGGCTGGGGGCGCGGGTCCGGGCGGCGCT	314
DB	241	TGTGGCTGCGCGTGACCCCGGAGCGCGGCTGGGGGCGCGGGTCCGGGCGGCGCT	300
QY	315	ACGAGGCTGCTGACAGAGTGGCGCGGCGGCGGCTTCAGGCGTGCAGCTGCTCAG	374
DB	301	ACGAGGCTGCTGACAGAGTGGCGCGGCGGCGGCTTCAGGCGTGCAGCTGCTCAG	360
QY	375	TGCTCTGCTACTGCGCGGCGGCGGCGGCGGCGGCTTCAGGCGTGCAGCTGCTC	434
DB	361	TGCTCTGCTACTGCGCGGCGGCGGCGGCGGCGGCTTCAGGCGTGCAGCTGCTC	420
QY	435	ACCCCTGAGTACCTGACACCGGCAAGCGGTGCTCGAGCTGCTGGCGGCTGCGAG	494
DB	421	ACCCCTGAGTACCTGACACCGGCAAGCGGTGCTCGAGCTGCTGGCGGCTGCGAG	480
QY	495	AGGACACCGCGGCTTGGGGAGTTCAGGCGGACCGCGGCGGCGGCTGCTGCGAG	554
DB	481	AGGACACCGCGGCTTGGGGAGTTCAGGCGGACCGCGGCGGCGGCTGCTGCGAG	540
QY	555	GCTCTGGAGGTGCAAGACGGCGGCTGCGAGGTGGGCTGCGCACAGGTGCTGCGG	614
DB	541	GCTCTGGAGGTGCAAGACGGCGGCTGCGAGGTGGGCTGCGCACAGGTGCTGCGG	600
QY	615	TCCGAGGCGCGGCTGCAACCGGTTGGTCCAGACTTGCACGTTCCGTTGCTTCCG	674
DB	601	TCCGAGGCGCGGCTGCAACCGGTTGGTCCAGACTTGCACGTTCCGTTGCTTCCG	660
QY	675	ACGGGAACCGCGGCGGCTTGGAGAGTGTACCTTCCTTTATTCCTGAAGCCCGG	734
DB	661	ACGGGAACCGCGGCGGCTTGGAGAGTGTACCTTCCTTTATTCCTGAAGCCCGG	720
QY	735	CAGTGTCTCAGCTGGTGCAGGCGGCGGCTGCGAGGTGGGCTGCGCACAGGTGCT	794
DB	721	CAGTGTCTCAGCTGGTGCAGGCGGCGGCTGCGAGGTGGGCTGCGCACAGGTGCT	780
QY	795	TTGCCATCGAAGACTGGATGCGACCGGTTAAACCCAGGTGACCCAGTCACTGGCA	854
DB	781	TTGCCATCGAAGACTGGATGCGACCGGTTAAACCCAGGTGACCCAGTCACTGGCA	840
QY	855	CACCTAAGGCTGCTCTTAAAGTCAACACCTCTTGCAITGGCCAGTGGAGAGATCT	914
DB	841	CACCTAAGGCTGCTCTTAAAGTCAACACCTCTTGCAITGGCCAGTGGAGAGATCT	900
QY	915	TTGATCATGAACCAACTATCATTTAGAGAGCTTTTACTCTTTGGCAATTATTTGG	974

DB	901	TTGATCATGAACCAACTATCATTTAGAGAGCTTTTACTCTTTGGCAATTATTTGG	960
QY	975	CTCCGAAATAGCTAAAGAAATCTGCAAAATCTCTGTGATTTGAGACAGGTACTGC	1034
DB	961	CTCCGAAATAGCTAAAGAAATCTGCAAAATCTCTGTGATTTGAGACAGGTACTGC	1020
QY	1035	GCAGGCCACCTATGCGCATAGCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCC	1094
DB	1021	GCAGGCCACCTATGCGCATAGCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCC	1080
QY	1095	CCCATCACCTGTGTACAGTGGCGAGGAGCTGCTCAAACTGACCTTATCTCTCTGC	1154
DB	1081	CCCATCACCTGTGTACAGTGGCGAGGAGCTGCTCAAACTGACCTTATCTCTCTGC	1140
QY	1155	TCACTGTGAGTCTGAGGAGAGTTGCGAGGCTGCAGGCGCGGGGCAATGCTCTTAC	1214
DB	1141	TCACTGTGAGTCTGAGGAGAGTTGCGAGGCTGCAGGCGCGGGGCAATGCTCTTAC	1200
QY	1215	GGGAAGACGAACTTGAGGCGAAGTGTGTTTCTCAAAAGGTAGAAATGCTCTTAC	1274
DB	1201	GGGAAGACGAACTTGAGGCGAAGTGTGTTTCTCAAAAGGTAGAAATGCTCTTAC	1260
QY	1275	AGCGATGAGAAATCTCTGGCTGCCATGTGGTTGATGCCAGCCCTCCAGAAAAGTCC	1334
DB	1261	AGCGATGAGAAATCTCTGGCTGCCATGTGGTTGATGCCAGCCCTCCAGAAAAGTCC	1320
QY	1335	TCAGACAGTATTAAGCCCTTAATCCAGAAATAGTTTGTAGTGAACCGTAGTTACT	1394
DB	1321	TCAGACAGTATTAAGCCCTTAATCCAGAAATAGTTTGTAGTGAACCGTAGTTACT	1380
QY	1395	GTGCCACGCTTAACATAGATTAGATGTTGTTGAAACATCTATCCACCATTTGTTAT	1454
DB	1381	GTGCCACGCTTAACATAGATTAGATGTTGTTGAAACATCTATCCACCATTTGTTAT	1440
QY	1455	AGTGTTCCTCAAAATTTCTGTCTTACAAAGATGTTGTGGCAGAAAACCTGGAGCA	1514
DB	1441	AGTGTTCCTCAAAATTTCTGTCTTACAAAGATGTTGTGGCAGAAAACCTGGAGCA	1500
QY	1515	TCCTAATTTTACTTCAGCCATCGTACCTCTTCTGATGATGAGCCGCTCATCAAAAG	1574
DB	1501	TCCTAATTTTACTTCAGCCATCGTACCTCTTCTGATGATGAGCCGCTCATCAAAAG	1560
QY	1575	TCCTCTCATCATGTTCAGTGAGGCGGCGAGTGTCTTCTCTGCGCATAGTAAACA	1634
DB	1561	TCCTCTCATCATGTTCAGTGAGGCGGCGAGTGTCTTCTCTGCGCATAGTAAACA	1620
QY	1635	TTTTCTTGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAAGACCTGT	1694
DB	1621	TTTTCTTGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAAGACCTGT	1680
QY	1695	CAGACAGCACAGGTGTACAGAAAGCAGCAGACAAGATCTTCAGATTCAGCAGGAG	1754
DB	1681	CAGACAGCACAGGTGTACAGAAAGCAGCAGACAAGATCTTCAGATTCAGCAGGAG	1740
QY	1755	CGGAGCCTCTGCTCTCTTACACTGGCATGCTGATGAGATCGTGACATGCCCATTTGC	1814
DB	1741	CGGAGCCTCTGCTCTCTTACACTGGCATGCTGATGAGATCGTGACATGCCCATTTGC	1800
QY	1815	TTCTTCCACATCTGGTTGCACTCGTCAATGATGGGCTCGTGCATCTCCCTCAGTCC	1874
DB	1801	TTCTTCCACATCTGGTTGCACTCGTCAATGATGGGCTCGTGCATCTCCCTCAGTCC	1860
QY	1875	TTCTAGTAGCCAAAGTGTTCCTCAGAGGCTGTCTATGTCTCTGGTCCCAAGGAGAC	1934
DB	1861	TTCTAGTAGCCAAAGTGTTCCTCAGAGGCTGTCTATGTCTCTGGTCCCAAGGAGAC	1918
QY	1935	TCCTGCAGAGCCATTTTGGGTAAGAACACATTAAGAAAGGCAATGATCTTGTCTG	1994
DB	1919	TCCTGCAGAGCCATTTTGGGTAAGAACACATTAAGAAAGGCAATGATCTTGTCTG	1978
QY	1995	AGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTAATAAGACATTCAGGCCA	2054
DB	1979	AGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTCATTAATAAGACATTCAGGCCA	2037

Db 182 GACGCCCCCGACCCCGCCCTGGCGGGCTGCTGGGGCCCCCGAGCGGAGCTACTCGCGT 241
Qy 257 TCGTGCCCGTGACCCCGGACCGCGCTGCGGGGCCCGGGTCCGGGCGCGGGCTGCAC 316
Db 242 TCGTGCCCGTGACCCCGGACCGCGCTGCGGGGCCCGGGTCCGGGCGCGGGCTGCAC 301
Qy 317 CAGCGCGCTGCTGACACCGTGGCGCGCGCGCCCTTCCAGGGTGCCAGTGTCCAGGCTG 376
Db 302 CAGCGCGCTGCTGACACCGTGGCGCGCGCGCCCTTCCAGGGTGCCAGTGTCCAGGCTG 361
Qy 377 CTCTGTACTGCG 436
Db 362 CTCTGTACTGCG 421
Qy 437 CCGCTGGATGACCCCTGACACCGCGGAAGCGCTGCTGAGCTGCTGGCGCCCTGCCAGGAG 496
Db 422 CCGCTGGATGACCCCTGACACCGCGGAAGCGCTGCTGAGCTGCTGGCGCCCTGCCAGGAG 481
Qy 497 GCACACG 556
Db 482 GCACACG 541
Qy 557 CTCTGGAGTGCACACCG 616
Db 542 CTCTGGAGTGCACACCG 601
Qy 617 CCGGAGCCCCCGTGCACCCGCTGGTGCAGACTTGCACAGTTCGCTGCTTCCGCGAC 676
Db 602 CCGGAGCCCCCGTGCACCCGCTGGTGCAGACTTGCACAGTTCGCTGCTTCCGCGAC 661
Qy 677 CCGGAGCG 736
Db 662 CCGGAGCG 721
Qy 737 GTGCTTGACCTGCTGCACAGTGCACCAACAGATCCAGAAAGAAAGTTCCAGGTTGTT 796
Db 722 GTGCTTGACCTGCTGCACAGTGCACCAACAGATCCAGAAAGAAAGTTCCAGGTTGTT 781
Qy 797 GCATCGAAGGACTGGATGCCCGGTAAAAACAAGTGCACCGTGCACCGTGCACCGTGC 856
Db 782 GCATCGAAGGACTGGATGCCCGGTAAAAACAAGTGCACCGTGCACCGTGCACCGTGC 841
Qy 857 CTTAAGCTGCTCTTAAAGTACACACCTCTGCAATGGCGGAGTGCAGTGGAGGATCTTT 916
Db 842 CTTAAGCTGCTCTTAAAGTACACACCTCTGCAATGGCGGAGTGCAGTGGAGGATCTTT 901
Qy 917 GATGATGAACCAACTATCAATGAGAGAGCTTTTACTCTTTGGGCAATTAATTTGGGCC 976
Db 902 GATGATGAACCAACTATCAATGAGAGAGCTTTTACTCTTTGGGCAATTAATTTGGGCC 961
Qy 977 TCCGAAATAGCTAAAGAAATCTCGCAAACTCTCTGTGANTGTAGACAGGTACTGGCACAG 1036
Db 962 TCCGAAATAGCTAAAGAAATCTCGCAAACTCTCTGTGANTGTAGACAGGTACTGGCACAG 1021
Qy 1037 ACGGCCACCTATGCCATAGCCATGAGGTGAGTGGGGTCTCCAGCAGCTGCCCGCCAGCC 1096
Db 1022 ACGGCCACCTATGCCATAGCCATGAGGTGAGTGGGGTCTCCAGCAGCTGCCCGCCAGCC 1081
Qy 1097 CATCACCTGTGTACAGTGGCGCAGAGGACCTGCTCAAACTTACCTGCTGCTGCTC 1156
Db 1082 CATCACCTGTGTACAGTGGCGCAGAGGACCTGCTCAAACTTACCTGCTGCTGCTC 1141
Qy 1157 ACTGTGAGTCTGTAGAGAGGTTGACAGGCTGACGGCGGGGCAATGAGAAAGACCGAG 1216
Db 1142 ACTGTGAGTCTGTAGAGAGGTTGACAGGCTGACGGCGGGGCAATGAGAAAGACCGAG 1201
Qy 1217 GAAGAGCAGAACTTGAAGCCACAGCTGTGTTTCGTCAAAAGTAGAATGCTTACACAG 1276
Db 1202 GAAGAGCAGAACTTGAAGCCACAGCTGTGTTTCGTCAAAAGTAGAATGCTTACACAG 1261
Qy 1277 CGGATGGAAATCCTGGCTGCCATGTGGTTGATGCCAGCCCTCCAGAGAAAGGTCCTG 1336
Db 1262 CGGATGGAAATCCTGGCTGCCATGTGGTTGATGCCAGCCCTCCAGAGAAAGGTCCTG 1321

Qy 1337 CAGACAGATTAAAGCCTAATCCAGAAATAGTTTATGTGAACCGTAGTTACTCTGGCCAGGT 1396
Db 1322 CAGACGCTAATTAAGCCTAATCCAGAAATAGTTTATGTGAACCGTAGTTACTCTGGCCAGGT 1381
Qy 1397 GCCAGCTCTAACTAGATTAGATTGTTTGAACATCTACATCCACCATTTGTTATGCAG 1456
Db 1382 GCCAGCTCTAACTAGATTAGATTGTTTGAACATCTACATCCACCATTTGTTATGCAG 1441
Qy 1457 TGTTCCCAAAATTTCTGTTCTACAAGCATGTTGTGGCAGAAAACCTGGAGACCAGCATC 1516
Db 1442 TGTTCCCAAAATTTCTGTTCTACAAGCATGTTGTGGCAGAAAACCTGGAGACCAGCATC 1501
Qy 1517 TTAATTTTACTTCAGCCATCGTACCCCTCTTCTGACTGATGGACCCGCTATCAAAAGTTC 1576
Db 1502 TTAATTTTACTTCAGCCATCGTACCCCTCTTCTGACTGATGGACCCGCTATCAAAAGTTC 1561
Qy 1577 CCTCTCATCATGTTCCAGTGAGAGGCCAGCGAATGCTTTCTTCTGGCATAGTAAACATT 1636
Db 1562 CCTCTCATCATGTTCCAGTGAGAGGCCAGCGAATGCTTTCTTCTGGCATAGTAAACATT 1621
Qy 1637 TTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAAGACCTGTTACTCA 1696
Db 1622 TTCTTGGAAACATATGTTTCACTTAATCACTACCAAAATATCTGGAAGACCTGTTACTCA 1681
Qy 1697 GACAGACAGGTGTACAGAGCAGCAGACAGATCTTCCAGATCAGAGGGAGACCCCG 1756
Db 1682 GACAGACAGGTGTACAGAGCAGCAGACAGATCTTCCAGATCAGAGGGAGACCCCG 1741
Qy 1757 GAGCCTCTGCTTCTCCTACACTGGCATGCTGATGAGATGCTGACATGCCACCATTTGGCTT 1816
Db 1742 GAGCCTCTGCTTCTCCTACACTGGCATGCTGATGAGATGCTGACATGCCACCATTTGGCTT 1801
Qy 1817 CTTCCACATCTGTTTGCACCTCGTCAATGATGGCTCGCTGCACTCCCTCAGTCCCAAAAT 1876
Db 1802 CTTCCACATCTGTTTGCACCTCGTCAATGATGGCTCGCTGCACTCCCTCAGTCCCAAAAT 1861
Qy 1877 CTAGTAGCAGAGTTCCTGCGAGAGCTGTCTATGTGCTGCTGGCTGCCCAAGGACACTC 1936
Db 1862 CTAG-AGCCAGTGTTCCTGCGAGAGCTGTCTATGTGCTGCTGGCTGCCCAAGGACACTC 1919
Qy 1937 CTGCGAGAGCCTTTTGGTGAAGAACCTTACAAGAGGCAATTCATTAAGACATTCAGCCAGATGC 1996
Db 1920 CTGCGAGAGCCTTTTGGTGAAGAACCTTACAAGAGGCAATTCATTAAGAGGCAATTCATTCGCTGAG 1979
Qy 1997 GCTCAGAGCCTTTTGTAGAGCTTCTGATGCTCAATTAAGACATTCAGCCAGATGC 2056
Db 1980 GCTCAGAGCCTTTTGTAGAGCTTCTGATGCTCAATTAAGACATTCAGCCAGATGC 2038
Qy 2057 TCCAACTGCAAAATATACCAACCTTCTGAAATTAATTTTGGCTTATTTATTTCTTTTC 2116
Db 2039 TCCAACTGCAAAATATACCAACCTTCTGAAATTAATTTTGGCTTATTTATTTCTTTTC 2098
Qy 2117 TTTTCTTCAAGAAATGCTCTGAAATGAATGACATTTCCATCTGAACCTGGATGCAT 2176
Db 2099 TTTTCTTCAAGAAATGCTCTGAAATGAATGACATTTCCAT-TGAACCTGGATGCAT 2156
Qy 2177 ATCAATTAAGCCTTCCAGTAAATTTATTAATTAATTAATTAATTAATTAATTTCTTCAGC 2236
Db 2157 TTAATTAAGCCTTCCAGTAAATTTATTAATTAATTAATTAATTAATTAATTTCTTCAGC 2216
Qy 2237 ATAGAGCTATGATTCATTAAT-AAAAGTGAGTCAAAAACCTAAATGCAATGTTTGT 2295
Db 2217 ATAGAGCTATGATTCATTAATTAATAAAGTGAGTCAAAAACCTAAATGCAATGTTTGT-T 2275
Qy 2296 GTGATTTTCAATACAAACCTTAATTTGCTGTTGTTAAATAGATTCAGTGCATCTTGA 2355
Db 2276 GTGATTTTCAATACAAACCTTAATTTGCTGTTGTTAAATAGATTCAGTGCATCTTGA 2334
Qy 2356 GTGGGATTTCTT-GGTAAATTAATCTT--GCACCTTGAATGCTCATGATTAATGATAA 2411
Db 2335 GTGGGATTTCTTGGGTAATTAATCTTTGCCCTTTGAAATGCTCTCTGATTAATGATAA 2394

```
QY 2412 TC-GCTTTGACATATCTTTAGACAGAAAAGTAGCTAGTGAGGGG-AAATTTATAG 2469
DB 2395 TCGGTTTGACATATCTTTAGACGGGAAAAGTAGCTAGTGAGGGGAAAATTTATAG 2454
QY 2470 CTTGTGTGACTTTAGGAGTAGCT 2493
DB 2455 GCTGTGTGACTTTAGGAGTAGCT 2478

RESULT 6
ADA38100
ID ADA38100 standard; DNA; 2856 BP.
XX AC ADA38100;
XX DT
XX 20-NOV-2003 (first entry)
XX DE
XX Thymidine kinase DNA, a target for antisense inhibition.
XX KW antisense; WNT-7B; N-acetylglucosaminyltransferase;
KW voltage-gated potassium channel; ion transport; Map3K8; thymidine kinase;
KW cell proliferation; H-Ras; small interfering RNA; siRNA; embryogenesis;
KW carcinogenesis; tumour progression; cell migration; matrix invasion;
KW cell differentiation; stress response; cytoskeletal; antiinflammatory;
KW cardiac arrhythmia; neurological disorder; epilepsy; interleukin 1b;
KW 1l-1b; ds.
XX OS
XX Unidentified.
XX FH
XX Key
XX 5'UTR
XX Location/Qualifiers
XX 1..27
XX /*tag= a
XX 28..1296
XX /*tag= b
XX /product= "Thymidine kinase protein"
XX 45..64
XX /*tag= c
XX /bound_moiety= "CG94235-01-TK-AS1"
XX /note= "Forms double stranded region with nucleotides 20-
XX 1 of sequence in {seqid:33}"
XX 66..85
XX /*tag= d
XX /bound_moiety= "CG94235-01-TK-AS2"
XX /note= "Forms double stranded region with nucleotides 20-
XX 1 of sequence in {seqid:34}"
XX 162..181
XX /*tag= e
XX /bound_moiety= "CG94235-01-TK-AS3"
XX /note= "Forms double stranded region with nucleotides 20-
XX 1 of sequence in {seqid:35}"
XX 659..678
XX /*tag= f
XX /bound_moiety= "CG94235-01-TK-AS4"
XX /note= "Forms double stranded region with nucleotides 20-
XX 1 of sequence in {seqid:36}"
XX 1297..2856
XX /*tag= g
XX /bound_moiety= "CG94235-01-TK-AS5"
XX /note= "Forms double stranded region with nucleotides 20-
XX 1 of sequence in {seqid:37}"
XX
XX WO2003070160-A2.
XX
XX 28-AUG-2003.
XX
XX 27-NOV-2002; 2002WO-US038188.
XX
XX 29-NOV-2001; 2001US-0334148P.
XX
XX 04-DEC-2001; 2001US-0336572P.
XX
XX 02-APR-2002; 2002US-00114153.
XX
XX 02-APR-2002; 2002US-00114270.
```

```
PR 01-MAY-2002; 2002US-00136826.
XX
XX (CURA-) CUPAGEN CORP.
XX
XX Ju J, Huang C, Zhong H, Simons JF, Taillon BE, Chant JS;
XX Peyman JA, Smithson G, Millet I;
XX WPI; 2003-697551/66.
XX
XX New oligonucleotides, useful in treatment and diagnosis of e.g. tumors,
XX inhibit expression of six specific genes, e.g. the oncogene WNT-7B, by
XX RNA interference.
XX
XX Claim 16; Page 48-49; 75pp; English.
XX
XX This invention relates to novel antisense oligonucleotides that modulate
XX the expression of WNT-7B, N-acetylglucosaminyltransferase, the voltage-
XX gated potassium channel, ion transport, Map3K8 or thymidine kinase.
XX Specifically, the invention describes inhibiting cell proliferation by
XX modulating the function of oncology targets: H-Ras, WNT-7B and
XX acetylglucosaminyltransferase. Small interfering RNA (siRNA) along with
XX the antisense compounds specifically hybridise to the target nucleic acid
XX molecules to inhibit gene expression. The Wnt proteins are secreted
XX ligands involved in embryogenesis and carcinogenesis, such that these
XX antisense oligos are useful for treating breast, gastric and colon
XX cancers. N-acetylglucosaminyltransferase are associated with tumour
XX progression, cell migration and matrix invasion, while Map3K8 regulates
XX cell differentiation and stress responses, such that antisense inhibitors
XX are cytostatic and antiinflammatory, and can be useful in cell
XX proliferative disorders. The voltage gated K channel maintains membrane
XX potential and modulates electrical excitability in neurons and can be
XX useful in the treatment of cardiac arrhythmias and neurological disorders
XX such as epilepsy. Thymidine kinase is important in DNA synthesis, and
XX antisense compounds can treat cell proliferation and modulate the
XX expression of interleukin 1b (1l-1b). Furthermore, antisense
XX oligonucleotides of the invention were designed to target H-ras and
XX interleukin 8 to inhibit their expression. This polynucleotide sequence
XX is the thymidine kinase DNA, a target of the invention.
XX
XX Sequence 2856 BP; 700 A; 712 C; 692 G; 752 T; 0 U; 0 Other;
XX
XX Query Match 88.5%; Score 2276; DB 9; Length 2856;
XX Best Local Similarity 98.9%; Pred. No. 0;
XX Matches 2388; Conservative 0; Mismatches 10; Indels 16; Gaps 9;
QY 83 GGGCGGCGGGGCTCGCGTGGGGCCATGGCTCCGCGCGCGCTTCGTCCTGGAGCTT 142
DB 1 GGGCGGCGGGGCTCGCGTGGGGCCATGGCTCCGCGCGCGCTTCGTCCTGGAGCTT 60
QY 143 CCGGACTGCACCCCTGGCTCACTTCGCCCTAGGCGCCGACGCCCCCGGCGAGCAGCC 202
DB 61 CCGGACTGCACCCCTGGCTCACTTCGCCCTAGGCGCCGACGCCCCCGGCGAGCAGCC 120
QY 203 CCGGACCCCGGCTGGGGGCGCTGTGGGGCCCCCGGAGGCGAGCTACTCGCTGGGTG 262
DB 121 CCGGACCCCGGCTGGGGGCGCTGTGGGGCCCCCGGAGGCGAGCTACTCGCTGGGTG 180
QY 263 CCGGTGACCCCGGAGCGCGGCTCGGGGGCCCGGCTCGGGGCGCGCTGCACACGCGC 322
DB 181 CCGGTGACCCCGGAGCGCGGCTCGGGGGCCCGGCTCGGGGCGCGGCTGCACACGCGC 240
QY 323 CTGCTGCACACAGCTGGCGCGCGCCCTTCACAGCGGTGCCAGTGTCTCAGCTGCTTGC 382
DB 241 CTGCTGCACACAGCTGGCGCGCGCCCTTCACAGCGGTGCCAGTGTCTCAGCTGCTTGC 300
QY 383 TACTGCCCGGGCGGCGAGCGCGGCGGCGACACAGGCTTCCTGTGCGCGACCCCGCTG 442
DB 301 TACTGCCCGGGCGGCGAGCGCGGCGGCGGCGACACAGGCTTCCTGTGCGCGACCCCGCTG 360
QY 443 GATGACCCCTGACACCCCGGCAAGCGCTGCTCGAGTGTCTGGGCGCGCTGCCAGGAGGACCA 502
DB 361 GATGACCCCTGACACCCCGGCGAGCGCTGCTCGAGTGTCTGGGCGCGCTGTCCAGGAGGACCA 420
```

QY 503 CGCCGCACTTGGGCGAGTTCAGGCGCGACCGCGGGCGACGCTGTGGCAGCGCCTCTGG 562
DB 421 CCCCCGCACTTGGGCGAGTTCAGGCGCGACCGCGGGCGACGCTGTGGCAGCGCCTCTGG 480
QY 563 GAGGTGCAAGACGGCAGCGGCTGCGAGTGGGCTGGCGACAGGCTGTGGCCCGTCCGGAG 622
DB 481 GAGGTGCAAGACGGCAGCGGCTGCGAGTGGGCTGGCGACAGGCTGTGGCCCGTCCGGAG 540
QY 623 CCCCCGCACTTGGGCGAGTTCAGGCGCGACCGCGGGCGACGCTGTGGCCCGTCCGGAG 682
DB 541 CCCCCGCACTTGGGCGAGTTCAGGCGCGACCGCGGGCGACGCTGTGGCCCGTCCGGAG 600
QY 683 GCGCCCGGGCGTTCGAGGAGTGTACCTCTTTATTTCTTCTGAAACCGGGCAGTGT 742
DB 601 GCGCCCGGGCGTTCGAGGAGTGTACCTCTTTATTTCTTCTGAAACCGGGCAGTGT 660
QY 743 GACCTGGTGGACGAGTCCCAAAACAGATCCAGAAAGGAAAGTTCAGGCTGTGGCCATC 802
DB 661 GACCTGGTGGACGAGTCCCAAAACAGATCCAGAAAGGAAAGTTCAGGCTGTGGCCATC 720
QY 803 GAAGGACTGGATGCCACG---GGTAAACACAGGTGACCCAGTCAAGTGCAGATTCACCT 859
DB 721 GAAGGACTGGATGCCACGCGGTGTAAACACAGGTGACCCAGTCAAGTGCAGATTCACCT 780
QY 860 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTCGCGAGTGGAGGAGATCTTGAT 919
DB 781 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTCGCGAGTGGAGGAGATCTTGAT 840
QY 920 GATGAACCAACTCATTTAGAAGAGCTTTTACTCTTTGGGCAATATATTTGGGCTCC 979
DB 841 GATGAACCAACTCATTTAGAAGAGCTTTTACTCTTTGGGCAATATATTTGGGCTCC 900
QY 980 GAAATAGCTAAAGATCTGCCAAATCTCTGTGATGTGTAGACAGGTACTGGCAGCAGC 1039
DB 901 GAAATAGCTAAAGATCTGCCAAATCTCTGTGATGTGTAGAC-----GGCAGCAGC 954
QY 1040 GCCACTATGCAATAGCACTCAGGTGAGTGGGGTCTCCAGCACTGCCCGCCAGCCCAT 1099
DB 955 GCCACTATGCAATAGCACTCAGGTGAGTGGGGTCTCCAGCACTGCCCGCCAGCCCAT 1014
QY 1100 CACCTGTGTACAGTGGCAGAGGACCTGCTCAACCTGACCTTATCTGTGCTCACT 1159
DB 1015 CACCTGTGTACAGTGGCAGAGGACCTGCTCAACCTGACCTTATCTGTGCTCACT 1074
QY 1160 GTGAGTCTGAGGAGGTGGCAGAGCTGCGAGGCGGGGATGAGAAAGCAGGGAA 1219
DB 1075 GTGAGTCTGAGGAGGTGGCAGAGCTGCGAGGCGGGGATGAGAAAGCAGGGAA 1134
QY 1220 GAAGCAGAACTTGAGGCCAACAGTGTGTTGTCGTCAGAAAGGTAGAAATGCTTACAGCG 1279
DB 1135 GAAGCAGAACTTGAGGCCAACAGTGTGTTGTCGTCAGAAAGGTAGAAATGCTTACAGCG 1194
QY 1280 ATGGAGATCTTGGTGGCAGTGTGATGCGAGCCCTCCAGAGAAAGGTCTCTGCAG 1339
DB 1195 ATGGAGATCTTGGTGGCAGTGTGATGCGAGCCCTCCAGAGAAAGGTCTCTGCAG 1254
QY 1340 ACAGTATTAAGCTTAATCCAGATAGTGTAGTGAACCGTAGTACTCTGCGCAGGTGCC 1399
DB 1255 ACGGTATTAAAGCTTAATCCAGATAGTGTAGTGAACCGTAGTACTCTGCGCAGGTGCC 1314
QY 1400 ACGTCTAAGTATAGATGTGTTGTAAGCACTTACATCCACCTTTGTTATGAGTGT 1459
DB 1315 ACGTCTAAGTATAGATGTGTTGTAAGCACTTACATCCACCTTTGTTATGAGTGT 1374
QY 1460 TCCCAATTTCTGTTCAAGCATGTTGTGGCAGAAACTGGAGACCGGCACTCTTA 1519
DB 1375 TCCCAATTTCTGTTCAAGCATGTTGTGGCAGAAACTGGAGACCGGCACTCTTA 1434
QY 1520 ATTTTACTTCAGCATCGTACCTCTTCTGATGAGCCCGTCAATCAAGAGTCCCT 1579
DB 1435 ATTTTACTTCAGCATCGTACCTCTTCTGATGAGCCCGTCAATCAAGAGTCCCT 1494
QY 1580 CTCATCATGTTCCAGTGAGAGGCCAGGATGCTTTCTTCTGGCATAGTAACATTTTC 1639

DB 1495 CTCATCATGTTCCAGTGAGAGGCCAGGANTGCTTTCTTCTGGCATAGTAACATTTTC 1554
QY 1640 TTGGAAACATATGTTTCACTTAATCACTCAAAATATCTGGAAGACCTGCTTACTCAGAC 1699
DB 1555 TTGGAAACATATGTTTCACTTAATCACTCAAAATATCTGGAAGACCTGCTTACTCAGAC 1614
QY 1700 AGCACAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGATCAGCAGGAGACCCGGAG 1759
DB 1615 AGCACAGGTGTACAGAAGCAGCAGCAAGATCTTCCAGATCAGCAGGAGACCCGGAG 1674
QY 1760 CCTCTGCTTCTCCTACACTGSCATGCTGATGAGATCTGACATGCCACATTTGGCTTCT 1819
DB 1675 CCTCTGCTTCTCCTACACTGSCATGCTGATGAGATCTGACATGCCACATTTGGCTTCT 1734
QY 1820 CCACATCTGGTTGCACTCGTCAATGATGGGCTCGCTGCACTCTCCCTCAGTCCCAATTTCTA 1879
DB 1735 CCACATCTGGTTGCACTCGTCAATGATGGGCTCGCTGCACTCTCCCTCAGTCCCAATTTCTA 1794
QY 1880 GTAGCCAAAGTGTCTCTGACAGGCTGTCTATGTCTCTGGCTGCCCAAGGAGACCTCTG 1939
DB 1795 G-AGCCAAAGTGTCTCTGACAGGCTGTCTATGTCTCTGGCTGCCCA-AGCACCTCTG 1852
QY 1940 CAGAGCCATTTTGGGTAAAGAACACTTACAAAGAAAGGCAATGATCTTGTGCTGAGGCT 1999
DB 1853 CAGAGCCATTTTGGGTAAAGAACACTTACAAAGAAAGGCAATGATCTTGTGCTGAGGCT 1912
QY 2000 CAGAGCCCTTTTGTAGGCTTCTGATGCTCAATTCATAAAGACATTCAGGCCAAGATGCTCC 2059
DB 1913 CAGAGCCCTTTTGTAGGCTTCTGA-GTCATATATAAGACATTCAGGCCAAGATGCTCC 1971
QY 2060 AACTGCAATATACCAACTCTCTGAAATATATTTGCTTAATTTATTTCTTTCTTT 2119
DB 1972 AACTGCAATATACCAACTCTCTGAAATATATTTGCTTAATTTATTTCTTTCTTT 2031
QY 2120 TTTTCTAAAGAAATGGCTCTGAAATAGATGCACATTTTCCATCTGAACCTGGATGCATATC 2179
DB 2032 TTTTCTAAAGTA-TGGCTCTGAAATAGATGCACATTTTCCAT-TGAACCTGGATGCATATC 2089
QY 2180 ATTTAGCCAAATCCAGTAAATTTATTTATATTAATCTATACATAATATGTTTCTCAGCATA 2239
DB 2090 ATTTAGCCAAATCCAGTAAATTTATTTATTAATCTATACATAATATGTTTCTCAGCATA 2149
QY 2240 GGAGCTATGATTCATTAATTAAGTGGAGTGCAGAAACGCTAAATGCAATGTTGTTGCT 2299
DB 2150 GGAGCTATGATTCATTAATTAAGTGGAGTGCAGAAACGCTAAATGCAATGTTGTTGCT 2209
QY 2300 ATTTTCAATACCAAACTTAATTTCTTGTAAATTAAGTTCAGTGGATCTTGGAGTGG 2359
DB 2210 ATTTTCAATACCAAACTTAATTTCTTGTAAATTAAGTAC-AGTGGATCTTGGAGTGG 2268
QY 2360 GATTTCTTGGTAAATTTATCTTGCATTTGAATGCTCATGATTCATATGAAATCGCTTTG 2419
DB 2269 GATTTCTTGGTAAATTTATCTTGCATTTGAATGCTCATGATTCATATGAAATCGCTTTG 2328
QY 2420 ACATATCTTTAGACAGAAAGTAGCTCAGTGGGGGGAATTTATAGAGCTTTGTGNC 2479
DB 2329 ACATATCTTTAGACAGAAAGTAGCTCAGTGGGGGGAATTTATAGAGCTTTGTGNC 2387
QY 2480 TTTAGGAGTAGTCT 2493
DB 2388 TTTAGGAGTAGTCT 2401

RESULT 7

ADC39135

ID ADC39135 standard; cDNA; 2856 BP.

XX ADC39135;

XX 18-DEC-2003 (first entry)

XX DE

Novel human NOVX polypeptide coding sequence SEQ ID NO: 77.

QY 803 GAAGGACTGGATGCCACG---GGTAAACCAACGGTGACCCAGTCACTGGGAGATTCACATT 859
DB |||||
721 GAAGGACTGGATGCCACGGGTGGTAAACCAACGGTGACCCAGTCACTGGGAGATTCACATT 780
QY |||||
DB |||||
860 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCAATGGCCAGTGGAGGAGATCTTTGAT 919
DB |||||
781 AAGGCTGTCTCTTAAAGTCAACACCTCTTGCAATGGCCAGTGGAGGAGATCTTTGAT 840
QY |||||
DB |||||
920 GATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGGCAATATATTGTGGCTCC 979
DB |||||
841 GATGAACCAACTATCATTAAGAGAGCTTTTACTCTTTGGGCAATATATTGTGGCTCC 900
QY |||||
DB |||||
980 GAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTGTAGACAGGTACTGGCAGACAG 1039
DB |||||
901 GAAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTGTAGAC-----GGCAGACAG 954
QY |||||
DB |||||
1040 GCCACCTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCGGCCCAT 1099
DB |||||
955 GGCACCTATGCCATAGCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCGGCCCAT 1014
QY |||||
DB |||||
1100 CACCCCTGTACCAAGTGGCCAGAGGACCTGCTCAAACTGACCTTATCTGCTGCTCACT 1159
DB |||||
1015 CACCCCTGTACCAAGTGGCCAGAGGACCTGCTCAAACTGACCTTATCTGCTGCTCACT 1074
QY |||||
DB |||||
1160 GTGAGTCTCTGAGGAGGTTGCAGAGGCTGCAGGGCCGGGGCATGGAGAAGACAGGGAA 1219
DB |||||
1075 GTGAGTCTCTGAGGAGGTTGCAGAGGCTGCAGGGCCGGGGCATGGAGAAGACAGGGAA 1134
QY |||||
DB |||||
1220 GAAGCAGAACTTTGAGGCCAACAGTGTGTTTGTCTCAAAAGGTAGAAATGCTTACACAGGG 1279
DB |||||
1135 GAAGCAGAACTTTGAGGCCAACAGTGTGTTTGTCTCAAAAGGTAGAAATGCTTACACAGGG 1194
QY |||||
DB |||||
1280 ATGGAGAACTCTGGTGGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGTCTCTGCAG 1339
DB |||||
1195 ATGGAGAACTCTGGTGGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGTCTCTGCAG 1254
QY |||||
DB |||||
1340 ACAGTATTAGCCTAATCCAGATAGTTTGTAGTGAACCGGTACTCTGCGCAGGTGCC 1399
DB |||||
1255 ACGGTATTAGCCTAATCCAGATAGTTTGTAGTGAACCGGTACTCTGCGCAGGTGCC 1314
QY |||||
DB |||||
1400 ACGTCTAAGTATAGATGTTGTTTGAACATCTACATCCACCTTTGTTATGCAATGCT 1459
DB |||||
1315 ACGTCTAAGTATAGATGTTGTTTGAACATCTACATCCACCTTTGTTATGCAATGCT 1374
QY |||||
DB |||||
1460 TCCCAAAATTTCTTTTACAAGCATGTTGTGGCAGAAATCTGGAGACCAGGCATCTTA 1519
DB |||||
1375 TCCCAAAATTTCTTTTACAAGCATGTTGTGGCAGAAATCTGGAGACCAGGCATCTTA 1434
QY |||||
DB |||||
1520 ATTTTACTTTCAGGCATCGTACCTCTTCTGACGTGATGGACCGGTATCAAAAGGTCCCT 1579
DB |||||
1435 ATTTTACTTTCAGGCATCGTACCTCTTCTGACGTGATGGACCGGTATCAAAAGGTCCCT 1494
QY |||||
DB |||||
1580 CTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTCTTCTGGCATAGTAAACATTTTC 1639
DB |||||
1495 CTCATCATGTTCCAGTGAGAGGCCAGCGATTGCTTTCTTCTGGCATAGTAAACATTTTC 1554
QY |||||
DB |||||
1640 TTGGAACATATGTTTCACTTAATCACTACCAATATCTGGAAGACCTGCTTACTCAGAC 1699
DB |||||
1555 TTGGAACATATGTTTCACTTAATCACTACCAATATCTGGAAGACCTGCTTACTCAGAC 1614
QY |||||
DB |||||
1700 AGCAGCAGTGTACAGAACAGCAGACAGATCTTCCAGATCAGAGGAGACCCCGGAG 1759
DB |||||
1615 AGCAGCAGTGTACAGAACAGCAGACAGATCTTCCAGATCAGAGGAGACCCCGGAG 1674
QY |||||
DB |||||
1760 CTTCTCTCTCTCTCTACACTGGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTT 1819
DB |||||
1675 CTTCTCTCTCTCTCTACACTGGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTT 1734
QY |||||
DB |||||
1820 CCAATCTCGTTGCACTCGTCAATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATCTTA 1879
DB |||||
1735 CCACATCTCGTTGCACTCGTCAATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATCTTA 1794

QY 1880 GTAGCCAAAGTGTCTCTGACAGAGGCTGTCTATGTGTCTGGCTGCCCAAGGACACTCCCTG 1939
DB |||||
1795 G-AGCCAAAGTGTCTCTGACAGAGGCTGTCTATGTGTCTGGCTGCCCA-AGACACTCCCTG 1852
QY |||||
DB |||||
1940 CAGAGCCATTTTGGTAAAGGAACACTTACAAGAGGACATTGATCTTGTCTGAGGCT 1999
DB |||||
1853 CAGAGCCATTTTGGTAAAGGAACACTTACAAGAGGACATTGATCTTGTCTGAGGCT 1912
QY |||||
DB |||||
2000 CAGAGCCCTTTGTATAGGCTTCTGATGTCAITTCATTAAGACATTCAGCAAGATCTCC 2059
DB |||||
1913 CAGAGCCCTTTGTATAGGCTTCTGAT-GTCAATATAAAGACATTCAGCAAGATCTCC 1971
QY |||||
DB |||||
2060 AACTGCAATATACCAACCTTCTCTGAATATATTTTGTCTTATTTATATTTCTTTCTTT 2119
DB |||||
1972 AACTGCAATATACCAACCTTCTCTGAATATATTTTGTCTTATTTATATTTCTTTCTTT 2031
QY |||||
DB |||||
2120 TTTTCTTAAAGAAATTTGCTCTGAATAGATGCAATTTTCCATCTGAATCTGGATGCATATC 2179
DB |||||
2032 TTTTCTTAAAGTA-TGGCTCTGAATAGATGCAATTTTCCAT-TGAACTTGGATGCATTTTC 2089
QY |||||
DB |||||
2180 ATTTAGCCAAATCCAGTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2239
DB |||||
2090 ATTTAGCCAAATCCAGTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2149
QY |||||
DB |||||
2240 GGAGCTATGATTCATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTGT 2299
DB |||||
2150 GGAGCTATGATTCATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTGT 2209
QY |||||
DB |||||
2300 ATTTTCAATACAAACTTAATTTGTTCTTGTAAATTAAGTTCAAAGTGGATCTTGGAGTGG 2359
DB |||||
2210 ATTTTCAATACAAACTTAATTTGTTCTTGTAAATTAAGTAC-AGTGGATCTTGGAGTGG 2268
QY |||||
DB |||||
2360 GATTTCTTGGTAAATTTATCTTGCATGTAATCTCTCATGATTTACATATGAATTCGCTTTG 2419
DB |||||
2269 GATTTCTTGGTAAATTTATCTTGCATGTAATCTCTCATGATTTACATATGAATTCGCTTTG 2328
QY |||||
DB |||||
2420 ACATATCTTTAGACAGAAAAGTACCTGAGTGGGGGAAATTTATAGACTTGTGTGAC 2479
DB |||||
2329 ACATATCTTTAGACAGAAAAGTACCTGAGTGGGGGAAATTTATAGACTTGTGTGAC 2387
QY |||||
DB |||||
2480 TTTAGGGAGTAGCT 2493
DB |||||
2388 TTTAGGGAGTAGCT 2401
RESULT 8
ADH42558
ID ADH42558 standard; DNA; 2856 BP.
XX ADH42558;
XX
XX 25-MAR-2004 (first entry)
DE Novel human nucleic acid NOV78a.
XX
XX ds; gene; cardiovascular; antiarteriosclerotic; hypotensive; cytostatic;
KW anorectic; antidiabetic; immunosuppressive; anti-HIV; neuroprotective;
KW neotropic; antiparkinsonian; antiasthmatic; antiinfertility;
KW cardiomyopathy; atherosclerosis; hypertension; cancer; obesity; diabetes;
KW AIDS; multiple sclerosis; graft-versus-host disease; Alzheimer's disease;
KW Parkinson's disease; asthma; fertility disorder; chromosome mapping;
XX tissue typing; preventive medicine; pharmacogenomic; vaccine.
OS Homo sapiens.
XX
XX WO2003102159-A2.
XX
XX 11-DEC-2003.
XX
XX 04-JUN-2003; 2003WO-US017573.
PF
XX
XX 04-JUN-2002; 2002US-0385490P.
PR
XX
XX 04-JUN-2002; 2002US-0385615P.

PR 04-JUN-2002; 2002US-0385755P.
PR 05-JUN-2002; 2002US-0386041P.
PR 06-JUN-2002; 2002US-0386355P.
PR 06-JUN-2002; 2002US-0386357P.
PR 06-JUN-2002; 2002US-0386447P.
PR 06-JUN-2002; 2002US-0386459P.
PR 06-JUN-2002; 2002US-0386465P.
PR 06-JUN-2002; 2002US-0386864P.
PR 07-JUN-2002; 2002US-0386701P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0387078P.
PR 07-JUN-2002; 2002US-0387081P.
PR 07-JUN-2002; 2002US-0387083P.
PR 10-JUN-2002; 2002US-0387429P.
PR 10-JUN-2002; 2002US-0387540P.
PR 10-JUN-2002; 2002US-0387866P.
PR 11-JUN-2002; 2002US-0387606P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387659P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 12-JUN-2002; 2002US-0388432P.
PR 12-JUN-2002; 2002US-0388479P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389604P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390144P.
PR 19-JUN-2002; 2002US-0390209P.
PR 25-JUN-2002; 2002US-0391726P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402268P.
PR 12-AUG-2002; 2002US-0402822P.
PR 13-AUG-2002; 2002US-0403458P.
PR 15-AUG-2002; 2002US-0403617P.
PR 15-AUG-2002; 2002US-0403732P.
PR 26-AUG-2002; 2002US-0406182P.
PR 12-SEP-2002; 2002US-0410085P.
PR 13-SEP-2002; 2002US-0410505P.
PR 23-SEP-2002; 2002US-0412955P.
PR 30-SEP-2002; 2002US-0415195P.
PR 23-OCT-2002; 2002US-0420627P.
PR 23-OCT-2002; 2002US-0420718P.
PR 24-OCT-2002; 2002US-0420852P.
PR 31-OCT-2002; 2002US-0422750P.
PR 01-NOV-2002; 2002US-0423095P.
PR 05-NOV-2002; 2002US-0423748P.

(CURA-) CURAGEN CORP.

XX PA Alsobrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;
XX PI Burgess CE, Casman SJ, Catterton E, Dhanabal M, Edinger SR;
PI Ellerman K, Ettenberg S, Gangalli EA, Gerlach VL, Gorman L;
PI Grosse WM, Gunther E, Guo X, Gusev VY, Herrmann JL, Ji W, Kekuda R;
PI Kiramtssov NV, Larochele WJ, Li L, Liang H, Low K, Macdougall JR;
PI MacLachlan T, Malyankar UN, Mcqueeney K, Mezick AJ, Miller CE;
PI Millet I, Padigaru M, Patturajan M, Peyman JA, Qian X, Rastelli L;
PI Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;
PI Spytek AR, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;
PI Wolenc KA, Zhong W, Zhong H;
XX WPI; 2004-053467/05.
DR P-PSDB; ADH42559.
XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
PT pharmacogenomics.

XX Claim 20; SEQ ID NO 1111; 1503pp; English.

XX The invention relates to 566 new isolated human polypeptides and their
CC encoding genes, sequences that are at least 95% identical to these or
CC sequences comprising one or more conservative substitutions in these. The
CC polypeptide, polynucleotide and antibodies against the polypeptides are
CC useful in diagnosing, treating or preventing NOVX-associated disorders,
CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
CC The nucleic acids are further used as hybridization probes, in chromosome
CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
CC polypeptides are also useful as vaccines. This sequence represents an
CC example of the nucleic acid sequence of the invention.

XX Sequence 2856 BP; 700 A; 712 C; 692 G; 752 T; 0 U; 0 Other;

Query Match 88.5%; Score 2276; DB 12; Length 2856;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 10; Indels 16; Gaps 9;

QY 83 GGGCGGCGGGGTCTGCGCTGGGGCCATGGCTCGCCGCTGCGGCTGCTGCTGGAGCTT 142
DB 1 GGGCGGCGGGGTCTGCGCTGGGGCCATGGCTCGCCGCTGCGGCTGCTGCTGGAGCTT 60
QY 143 CCGACTGCACCTTGGCTCACTTGCCTTAGGCGCCGACGCCCCCGGCGAGCAGCGCC 202
DB 61 CCGACTGCACCTTGGCTCACTTGCCTTAGGCGCCGACGCCCCCGGCGAGCAGCGCC 120
QY 203 CCGACCCCGCGCTGGCGGCTGTTGGGGCCCCCGAGGAGGAGTACTCGCTGGCTG 262
DB 121 CCGACCCCGCGCTGGCGGCTGCTGGGGCCCCCGAGGAGGAGTACTCGCTGGCTG 180
QY 263 CCGTGACCCCGGACCGCGGCTGCGGGGCCCGGGTCCGGCGCGCGGCTGCACACGCGC 322
DB 181 CCGTGACCCCGGACCGCGGCTGCGGGGCCCGGGTCCGGCGCGCGGCTGCACACGCGC 240
QY 323 CTGCTGCACACAGTGGCGCGCGCCCTTCCAGCGTGCAGTGTCTCAGCTGTCTGTC 382
DB 241 CTGCTGCACACAGTGGCGCGCGCCCTTCCAGCGTGCAGTGTCTCAGCTGTCTGTC 300
QY 383 TACTGCCCGGGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 442
DB 301 TACTGCCCGGGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 443 GATGACCTGACACCGCGGCAAGCGCTGCTGAGCTGTGGGGCGCGCTGCCAGGAGGACCA 502
DB 361 GATGACCTGACACCGCGGCAAGCGCTGCTGAGCTGTGGGGCGCGCTGCCAGGAGGACCA 420
QY 503 CGCCCGCAGCTTGGCGGAGTTTCGAGGGCGACCCCGCGCGCGCGCGCGCGCGCGCGCG 562
DB 421 CGCCCGCAGCTTGGCGGAGTTTCGAGGGCGACCCCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 563 GAGTGCAAGACCGGAGCGCGCTGCAGTGGGCTGGCGACAGGTGCTGCCGCTCCCGGAG 622
DB 481 GAGTGCAAGACCGGAGCGCGCTGCAGTGGGCTGGCGACAGGTGCTGCCGCTCCCGGAG 540
QY 623 CCCCCTGTCACCCCGGTGGTGCAGACTTGCACAGTTCCGTTGGTCTTCCCGGACCGGAA 682
DB 541 CCCCCTGTCACCCCGGTGGTGGTGCAGACTTGCACAGTTCCGTTGGTCTTCCCGGACCGGAA 600
QY 683 GCGCCCGCGCGGTTCGAGGAGTGTACCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 742
DB 601 GCGCCCGCGCGGTTCGAGGAGTGTACCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
QY 743 GACCTGTCACAGTGCCTCCCAACACAGATCCAGAAAGAAAGTTCCAGGTGTTGTCATC 802
DB 661 GACCTGTCACAGTGCCTCCCAACACAGATCCAGAAAGAAAGTTCCAGGTGTTGTCATC 720

FT	misc_feature	629..708		QY	83	GGGCGCGCGGGTCTGGCGCTGGGGCCATGGCTCCGCGCTGCGCTTCTCTGGAGCTT	142
FT	/*tag= c						
FT	/note= "Target region for antisense oligonucleotides.Claimed in claim 16"			Db	1	GGCGCGCGGGTCTGGCTGGGGCCATGGCTCCGCGCGCGCTTCTCTGGAGCTT	60
FT	1165..1286						
FT	/*tag= d			QY	143	CCCGACTGCACCTGGCTACCTTGGCGCTTAGCGCGCGAGCGCGCGCGCGAGCAGCAGCGC	202
FT	1165..1286			Db	61	CCCGACTGCACCTGGCTACCTTGGCGCGCGAGCGCGCGCGCGAGCAGCAGCGC	120
FT	/note= "Target region for antisense oligonucleotides.Claimed in claim 16"			QY	203	CCCGACCCCGCGCTGGCGCGCTGTGGGCGCGCGCGAGCGAGCGAGCTACTCGCTGTGCGTG	262
FT				Db	121	CCCGACCCCGCGCTGGCGCGCTGTGGGCGCGCGCGAGCGAGCGAGCTACTCGCTGTGCGTG	180
PN	US2003176385-A1.			QY	263	CCCGTGACCCCGAGCGCGCGCTGCGGGCGCGCGGTCCGGCGCGCGGTGACACAGCGC	322
XX	18-SEP-2003.			Db	181	CCCGTGACCCCGAGCGCGCGCTGCGGGCGCGCGGTCCGGCGCGCGGTGACACAGCGC	240
XX	27-NOV-2002; 2002US-00305810.			QY	323	CTGCTGACACAGTGGCGCGCGCGCTTCCAGCGGTGCGAGCTGTCTAGGGTGTCTGTC	382
PR	15-FEB-2000; 2000US-0182637P.			Db	241	CTGCTGACACAGTGGCGCGCGCGCTTCCAGCGGTGCGAGCTGTCTAGGGTGTCTGTC	300
PR	29-MAR-2000; 2000US-0192838P.			QY	383	TACTGCCCGGGCGCGCGCGCGCGCGAGCGCGCGAGCGCTTCTGCTGGCGCGACCCCTG	442
PR	03-APR-2000; 2000US-0194256P.			Db	301	TACTGCCCGGGCGCGCGCGCGCGCGAGCGCGCGAGCGCTTCTGCTGGCGCGACCCCTG	360
PR	26-JUL-2000; 2000US-00625634.			QY	443	GATGACCTGACACACCGCGCAAGCGTGTCTGAGCTCTGGGCGCGCTGCCAGAGGACCA	502
PR	29-NOV-2001; 2001US-0334148P.			Db	361	GATGACCTGACACACCGCGCAAGCGTGTCTGAGCTCTGGGCGCGCTGCCAGAGGACCA	420
XX	04-DEC-2001; 2001US-0336572P.			QY	503	CGCCCGCACTTGGCGAGTTCAGGGCGAGCGCGCGCGCGCGAGCGCTTCTGCTGGCGAGCGCGCTG	562
XX	(JULJ/) JU J.			Db	421	CGCCCGCACTTGGCGAGTTCAGGGCGAGCGCGCGCGCGCGAGCGCTTCTGCTGGCGAGCGCGCTG	480
PA	(HUAN/) HUANG C.			QY	563	GAGGTGCAAGACCGCGAGCGCGTGCAGGTGGGCTGCGCACAGCTCGTGGCGCGCTGCCAG	622
PA	(ZHON/) ZHONG H.			Db	481	GAGGTGCAAGACCGCGAGCGCGTGCAGGTGGGCTGCGCACAGCTCGTGGCGCGCTGCCAG	540
PA	(SIMO/) SIMONS J F.			QY	623	CCCGCGCTGACACCGCGTGGTGGCGAGCTTCCAGTTCCTGGTCTTCCCGGACCGGAA	682
PA	(TAIL/) TAILLON B E.			Db	541	CCCGCGCTGACACCGCGTGGTGGCGAGCTTCCAGTTCCTGGTCTTCCCGGACCGGAA	600
PA	(CHAN/) CHANT J S.			QY	683	CGCGCGCGCGCGTGGTGGAGGAGTACCTCTTTATTCCTGAAGCGCGCGGAGTCTT	742
PA	(PEYM/) PEYMAN J A.			Db	601	CGCGCGCGCGCGTGGTGGAGGAGTACCTCTTTATTCCTGAAGCGCGCGGAGTCTT	660
PA	(SMIT/) SMITHSON G.			QY	743	GACTGTGTACACAGTCCCAAAACAGATCCAGAAAGAAAGTTCCAGTGTGTGCCATC	802
PA	(MILL/) MILLET I.			Db	661	GACTGTGTACACAGTCCCAAAACAGATCCAGAAAGAAAGTTCCAGTGTGTGCCATC	720
XX	Ju J, Huang C, Zhong H, Simons JF, Taillon BE, Chant JS; Peyman JA, Smithson G, Millet I; WPI; 2003-898588/82.			QY	803	GAAGGACTGGATGCCACG---GGTAAACCAACCGGTGCTGACCCAGTCCAGTGGCAGATTCATT	859
XX	New antisense oligonucleotide for modulating expression of WNT-7B, N-acetylglucosaminyltransferase, voltage-gated K channel, ion transport, Map3K8 or thymidine kinase, or for treating cancer, infection or inflammation.			Db	721	GAAGGACTGGATGCCACG---GGTAAACCAACCGGTGCTGACCCAGTCCAGTGGCAGATTCATT	780
XX	Claim 16; SEQ ID NO 6; 43pp; English.			QY	860	AAGCTGTCTCTTAAAGTACCACCTCTTGAATGGCGCGCGAGTGGAGAGATCTTTGAT	919
XX	The invention relates to an oligonucleotide (antisense) 20-50 or 10-50 nucleotides in length that is targeted to regions of the cDNAs appearing as ADM10596-ADM10601 (human WNT-7B, N-acetylglucosaminyltransferase, voltage-gated K channel, ion transport channel, Map3K8 (mitogen activated protein kinase kinase 8) or thymidine kinase. Also included are methods of inhibiting the expression of WNT-7B, N-acetylglucosaminyltransferase, voltage-gated K channel, ion transport, Map3K8 or thymidine kinase in a cell, comprising contacting the cell with the oligonucleotide cited above, a method of inhibiting cell proliferation (comprising contacting a cell with the oligonucleotide cited above) and a method of increasing the production of IL-1b (interleukin 1b) in a cell (comprising contacting a cell with the oligonucleotide cited above). Also disclosed are antisense oligonucleotides for H-ras and interleukin-8. The antisense oligonucleotide is useful in modulating the expression of WNT-7B, N-acetylglucosaminyltransferase, voltage-gated K channel, ion transport channel, Map3K8 or thymidine kinase to treat diseases associated with their expression, such as cell proliferative disorders, infection or inflammation. In addition, the composition is used for diagnostics, prophylaxis, or as research reagents or kits. The present sequence is the cDNA from one of the genes listed above, a target for the antisense oligonucleotides of the invention.			Db	781	AAGCTGTCTCTTAAAGTACCACCTCTTGAATGGCGCGCGAGTGGAGAGATCTTTGAT	840
XX	Sequence 2856 BP; 700 A; 713 C; 691 G; 752 T; 0 U; 0 Other;			QY	920	GATGAACCAACTATCATTAGAAGCTTTTACTCTTTGGGCAATTATTTGGGCTCC	979
XX	Query Match 88.5%; Score 2274.4; DB 11; Length 2856;			Db	841	GATGAACCAACTATCATTAGAAGCTTTTACTCTTTGGGCAATTATTTGGGCTCC	900
XX	Best Local Similarity 98.9%; Pred. NO. 0;			QY	980	GAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTAGACAGGTACTGGCAGACAGC	1039
XX	Matches 2387; Conservative 0; Mismatches 11; Indels 16; Gaps 9;			Db	901	GAATAGCTAAAGAAATCTGCCAAATCTCTGTGATTTAGACAGGTACTGGCAGACAGC	954
XX				QY	1040	GCCACCTATGCCATAGCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCGCAGCCCAT	1099
XX				Db	955	GCCACCTATGCCATAGCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCGCAGCCCAT	1014
XX				QY	1100	CACCTGTGTACCAAGTGGCCAGAGCACTGTCTAAACCTGACCTTATCTCTGCTGCTACT	1159
XX				Db	1015	CACCTGTGTACCAAGTGGCCAGAGCACTGTCTAAACCTGACCTTATCTCTGCTGCTACT	1074
XX				QY	1160	GTGAGTCTGTAGGAGAGTTCCAGAGGTGCGAGGCGCGCGCGCGAGAGACAGGAA	1219

```
Db 1075 GTGAGTCTTGAGGAGAGGTTGACAGAGCTGACGGCGGGGNCATGAGAGACCGAGGAA 1134
QY 1220 GAAGCAGAACTTGAGCCCAACAGTGTGTTTCGTCAAAAGGTAGAAATGCTCTACACGCG 1279
Db 1135 GAAGCAGAACTTGAGCCCAACAGTGTGTTTCGTCAAAAGGTAGAAATGCTCTACACGCG 1194
QY 1280 ATGGAGAAATCTGGCTGCCATGTGGTTGATGCCAGGCCCTCCAGAGAAAGGTCCTGGAG 1339
Db 1195 ATGGAGAAATCTGGCTGCCATGTGGTTGATGCCAGGCCCTCCAGAGAAAGGTCCTGGAG 1254
QY 1340 ACAGTATTAAAGCCCTAATCCAGAAATAGTTTTAGTGAACCGGTAGTACTCTGGCCAGGTGCC 1399
Db 1255 ACGGTATTAAAGCCCTAATCCAGAAATAGTTTTAGTGAACCGGTAGTACTCTGGCCAGGTGCC 1314
QY 1400 AGCTCTAACTAGANTAGATGTTGTTTGAACAPCTACATCCACCAATTTGTTATGCAGTGT 1459
Db 1315 AGCTCTAACTAGANTAGATGTTGTTTGAACAPCTACATCCACCAATTTGTTATGCAGTGT 1374
QY 1460 TCCCAAAATTTCTGTTCTACAGCATGTTGTGTGGCAGAAACTGGAGACCGGCACTCTTA 1519
Db 1375 TCCCAAAATTTCTGTTCTACAGCATGTTGTGTGGCAGAAACTGGAGACCGGCACTCTTA 1434
QY 1520 ATTTTACTTTCAGCCATCGTACCCCTCTCTGACTGATGGACCCGTCATCAAAAGGTCCTCT 1579
Db 1435 ATTTTACTTTCAGCCATCGTACCCCTCTCTGACTGATGGACCCGTCATCAAAAGGTCCTCT 1494
QY 1580 CTCATCATGTTCCAGTGAGAGGCCAGCGATGCTTCTTCTGGCAGATGTAACCAATTTTC 1639
Db 1495 CTCATCATGTTCCAGTGAGAGGCCAGCGATGCTTCTTCTGGCAGATGTAACCAATTTTC 1554
QY 1640 TTGGACATATGTTTCACTTAATCACTACCAATATCTGGAGACCTGTCTTACTCAGAC 1699
Db 1555 TTGGACATATGTTTCACTTAATCACTACCAATATCTGGAGACCTGTCTTACTCAGAC 1614
QY 1700 AGCACAGGTGTACAGAGCAGACAGAAAGATCTTCCAGATCAGCAGGGAGACCCCGGAG 1759
Db 1615 AGCACAGGTGTACAGAGCAGACAGAAAGATCTTCCAGATCAGCAGGGAGACCCCGGAG 1674
QY 1760 CTTCTGCTTCTCTACCTACCTGCGATGCTGATGAGATCGTGACATGCGCCACATTTGCTTT 1819
Db 1675 CTTCTGCTTCTCTACCTACCTGCGATGCTGATGAGATCGTGACATGCGCCACATTTGCTTT 1734
QY 1820 CCACATCTGTTGCACTGCTCATGATGGGCTGCTGCTCATCTCCCTCAGTCCCAATTTCTA 1879
Db 1735 CCACATCTGTTGCACTGCTCATGATGGGCTGCTGCTCATCTCCCTCAGTCCCAATTTCTA 1794
QY 1880 GTAGCCAAAGTGTCTTGCAGAGGCTGTATGTGCTGCTGCCCAAGGACACCTCTG 1939
Db 1795 G-AGCCAAAGTGTCTTGCAGAGGCTGTATGTGCTGCTGCCCA-GGACATCTCTG 1852
QY 1940 CAGAGCCATTTTGGGTAAAGAACACTTACAAAGAGGCAATGATTTGTGCTGAGGCT 1999
Db 1853 CAGAGCCATTTTGGGTAAAGAACACTTACAAAGAGGCAATGATTTGTGCTGAGGCT 1912
QY 2000 CAGAGCCCTTTTGAATAGGCTTCTGATGTCATTCATAAGACATTCAGGCCAGATGCTCC 2059
Db 1913 CAGAGCCCTTTTGAATAGGCTTCTGATGTCATTCATAAGACATTCAGGCCAGATGCTCC 1971
QY 2060 AACTGCAAAATATACCAACCTTCTGAAATATATTTTGTGTTATTTATATTTCTTTCTTT 2119
Db 1972 AACTGCAAAATATACCAACCTTCTGAAATATATTTTGTGTTATTTATATTTCTTTCTTT 2031
QY 2120 TTTTCTAAAGAAATGCTCTGTAATAGAAAGCAATTTTCCATCTGAATGCGATGCAATATC 2179
Db 2032 TTTTCTAAAGAAATGCTCTGTAATAGAAAGCAATTTTCCAT-TGAATGCGATGCAATTTTC 2089
QY 2180 ATTTAGCCAAATCCAGTAAATTTATTTATATATATATATATATATATGTTTCTCAGCAT 2239
Db 2090 ATTTAGCCAAATCCAGTAAATTTATTTATATATATATATATATATATGTTTCTCAGCAT 2149
QY 2240 GGAGCTATGATTCATTAATAATAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGCT 2299
```

```
Db 2150 GGAGCTATGATTCATTAATTAATAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGT 2209
QY 2300 ATTTTCATTACACAAACTTAATTTGTTCTTTAAATAAGTTCAAAGTGGATCTTTGGAGTGG 2359
Db 2210 ATTTTCATTACACAAACTTAATTTGTTCTTTAAATAAGTAC-AGTGGATCTTTGGAGTGG 2268
QY 2360 GATTTCTTGGTAAATTAATCTTGCACCTTGAATGTCTCATCATGATACATATGAAATCGCTTTG 2419
Db 2269 GATTTCTTGGTAAATTAATCTTGCACCTTGAATGTCTCATCATGATACATATGAAATCGCTTTG 2328
QY 2420 ACATATCTTTAGACAGAAAAAAGTAGCTCAGTGGAGGGGAAATTTATAGAGCTTTGTGTGAC 2479
Db 2329 ACATATCTTTAGACAGAAAAAAGTAGCTCAGTGGAGGGGAAATTTATAGAGC-TGTTGTGAC 2387
QY 2480 TTTAGGAGTAGCT 2493
Db 2388 TTTAGGAGTAGCT 2401

RESULT 10
ABX63720
ID ABX63720 standard; cDNA; 2645 BP.
XX
AC ABX63720;
XX
DT 26-FEB-2003 (first entry)
XX
DE Human cDNA #720 differentially expressed in activated vascular tissue.
XX
KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
KW gene therapy; vascular disease; cancer; coronary; artery disease;
KW hypertension; diabetes; pre-eclampsia; restenosis;
KW ischaemia-reperfusion injury; stroke.
XX
OS Homo sapiens.
XX
PN US2002137081-A1.
XX
PD 26-SEP-2002.
XX
PF 08-JAN-2002; 2002US-00044090.
XX
PR 28-JUL-2000; 2000US-0222469P.
PR 08-JAN-2001; 2001US-0260483P.
XX
PA (BAND/) BANDMAN O.
XX
PI Bandman O;
XX
WPI; 2003-110597/10.
XX
Combination for diagnosing, staging, treating, or monitoring the
progression of treatment of a vascular disease, e.g. atherosclerosis,
comprises several cDNAs that are differentially expressed in activated
vascular tissue.
Claim 1; Page; 18pp; English.
XX
This invention relates to a combination comprising several cDNAs that are
differentially expressed in activated vascular tissue. The invention also
discloses a high throughput method for detecting differentially expressed
cDNAs in a sample. The cDNAs of the invention may have
antiartherosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
gynaecological; vasotropic and cerebroprotective activities and may be
used in gene therapy. The cDNAs of the invention may be used in a high-
throughput methods for detecting differential expression of one or more
cDNAs in a sample, or screening several molecules or compounds to
identify a molecule or compound that specifically binds a cDNA of the
invention. A protein encoded by the cDNA may be used to screen several
molecules or compounds to identify a ligand that specifically binds to
the protein, or to produce or purify an antibody to the protein that can
be used to detect a protein in a sample or purify a natural or
```

CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of treatment
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion
CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
CC genetic or gene expression analysis of several new nucleic acid
CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
CC diagnosing pre-pathologic disorders, and chronic or acute diseases
CC associated with abnormalities in the expression, amount or distribution
CC of the protein. The present sequence represents a cDNA of the invention
CC that is differentially expressed in activated vascular tissue. Note: The
CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC <http://seqdata.uspto.gov/sequence.html?docID=20020137081>
XX
SQ

Sequence 2645 BP; 671 A; 628 C; 622 G; 724 T; 0 U; 0 Other;

Query Match		82.7%;	Score 2125.2;	DB 8;	Length 2645;
Best Local Similarity		99.3%;	Pred. No. 0;		
Matches 2207;		Conservative	0;	Mismatches	8;
				Indels	7;
				Gaps	7;
QY	272	CCGGACGCCGGCTGCGGGGCCCGGGTCCAGCGGTGCAGCTGCTCAGGCTGCTCTACTGCCCG	331		
DB	1	CCGGACGCCGGCTGCGGGGCCCGGGTCCAGCGGTGCAGCTGCTCAGGCTGCTCTACTGCCCG	60		
QY	332	CAGCTGCGCGCGCGCCCTTCCAGCGGTGCAGCTGCTCAGGCTGCTCTACTGCCCG	391		
DB	61	CAGCTGCGCGCGCGCCCTTCCAGCGGTGCAGCTGCTCAGGCTGCTCTACTGCCCG	120		
QY	392	GGCGGCGCAGCGCGCGCGCACAGCAAGGCTTCTGTGCGCGCACCCCTGGATGACCTT	451		
DB	121	GGCGGCGCAGCGCGCGCGCACAGCAAGGCTTCTGTGCGCGCACCCCTGGATGACCTT	180		
QY	452	GACACCCCGCAAGCGTGTCTGAGCTGCTGGGCGCGCTGCCAGGAGCACACCGCCGCGAC	511		
DB	181	GACACCCCGCAAGCGTGTCTGAGCTGCTGGGCGCGCTGCCAGGAGCACACCGCCGCGAC	240		
QY	512	TTGGGCGAGTTCAGGCG	571		
DB	241	TTGGGCGAGTTCAGGCG	300		
QY	572	GAGGCGAGGCGGTGCAAGTGGGCTGCGCACAGTCTGTGCGCGCGCGCGCGCGCGCGCG	631		
DB	301	GAGGCGAGGCGGTGCAAGTGGGCTGCGCACAGTCTGTGCGCGCGCGCGCGCGCGCGCG	360		
QY	632	CACCGGTGTGCGCAGTTGCCAGTTCCGAGTTCCGAGTTCCGAGTTCCGAGTTCCGAGTT	691		
DB	361	CACCGGTGTGCGCAGTTGCCAGTTCCGAGTTCCGAGTTCCGAGTTCCGAGTTCCGAGTT	420		
QY	692	CGCGTTTGGAGGAGTGTACCTCTTTATTCCTGAAGCCCGGCGAGTGTGACCTGCTC	751		
DB	421	CGCGTTTGGAGGAGTGTACCTCTTTATTCCTGAAGCCCGGCGAGTGTGACCTGCTC	480		
QY	752	GACAGTGCCTCAAAACAGATCCAGAAAGGAAAGTTCCAGTGTGTGCCATCGAAGGACTG	811		
DB	481	GACAGTGCCTCAAAACAGATCCAGAAAGGAAAGTTCCAGTGTGTGCCATCGAAGGACTG	540		
QY	812	GATGCCACGGTAAACACAGGTGACCCAGTCCAGTGGCAGATTCACTTAAGGCTGTCTC	871		
DB	541	GATGCCACGGTAAACACAGGTGACCCAGTCCAGTGGCAGATTCACTTAAGGCTGTCTC	600		
QY	872	TTAAGTCAACCCCTCTTGATTTGGCGAGTGGAGGAGATCTTTGATGATGAACCAACT	931		
DB	601	TTAAGTCAACCCCTCTTGATTTGGCGAGTGGAGGAGATCTTTGATGATGAACCAACT	660		
QY	932	ATCATTAGAAGAGCTTTTACTCTTTGGGCAATTAATTTGGGCTCCGAATAGCTAAA	991		
DB	561	ATCATTAGAAGAGCTTTTACTCTTTGGGCAATTAATTTGGGCTCCGAATAGCTAAA	720		
QY	992	GAATCTGCAATCTCTGTGATTTGTAGACAGGTACTGGCAGCAGCGCCACCTATGCC	1051		
DB	721	GAATCTGCAATCTCTGTGATTTGTAGACAGGTACTGGCAGCAGCGCCACCTATGCC	780		

QY	1052	ATAGCCACTGAGGTGAGTGGGGTCTCCAGACACCTGCCCCAGCCCATCACCTGTGTATC	1111		
DB	781	ATAGCCACTGAGGTGAGTGGGGTCTCCAGACACCTGCCCCAGCCCATCACCTGTGTATC	840		
QY	1112	CAGTGGCCAGAGGACTGCTCAAACTGACCTTATCTCTGCTGCTCAGTCTGAGTCTGAG	1171		
DB	841	CAGTGGCCAGAGGACTGCTCAAACTGACCTTATCTCTGCTGCTCAGTCTGAGTCTGAG	900		
QY	1172	GAGAGTTCAGAGGCTGCGAGGCGCGGGCATGGAGAACACAGGAGGAGAGCAGAACCTT	1231		
DB	901	GAGAGTTCAGAGGCTGCGAGGCGCGGGCATGGAGAACACAGGAGGAGAGCAGAACCTT	960		
QY	1232	GAGGCAACAGTGTGTTTCCTCAAAAGGTAGAAATGCTCTACAGGCGATGGAGATCTCT	1291		
DB	961	GAGGCAACAGTGTGTTTCCTCAAAAGGTAGAAATGCTCTACAGGCGATGGAGATCTCT	1020		
QY	1292	GGCTGCTATGTTGATGCCAGCCCTCCAGAGAAAGGTCTCTGAGACAGTATTAAAGC	1351		
DB	1021	GGCTGCTATGTTGATGCCAGCCCTCCAGAGAAAGGTCTCTGAGACAGTATTAAAGC	1080		
QY	1352	CTAATCCAGAAATAGTTTATGTAACCGTAGTTACTCTGGCCAGGTGCCAGTCTAACTAG	1411		
DB	1081	CTAATCCAGAAATAGTTTATGTAACCGTAGTTACTCTGGCCAGGTGCCAGTCTAACTAG	1140		
QY	1412	ATTAGATGTTGTTGAACATCTACATCCACATTTGTTATGCAAGTGTCCCAAAATTTCT	1471		
DB	1141	ATTAGATGTTGTTGAACATCTACATCCACATTTGTTATGCAAGTGTCCCAAAATTTCT	1200		
QY	1472	GTTCTAAGATGTTGTTGCGCAGAAACTGAGACACAGGATCTTAATTTTACTTTCTAG	1531		
DB	1201	GTTCTAAGATGTTGTTGCGCAGAAACTGAGACACAGGATCTTAATTTTACTTTCTAG	1260		
QY	1532	CCATCTGACCTCTTCTGACTGATGACCGCTCATCAAAAGGTCCCTCTCATCTGTTTC	1591		
DB	1261	CCATCTGACCTCTTCTGACTGATGACCGCTCATCAAAAGGTCCCTCTCATCTGTTTC	1320		
QY	1592	CAGTGAAGGCCAGCAGTGTCTTTCTTCTGGCATAGTAAACAATTTCTTGGACATATG	1651		
DB	1321	CAGTGAAGGCCAGCAGTGTCTTTCTTCTGGCATAGTAAACAATTTCTTGGACATATG	1380		
QY	1652	TTTCACTTAATCACTACCAATATCTGGAGACCTGTCTTACTCAGACACACACAGTGT	1711		
DB	1381	TTTCACTTAATCACTACCAATATCTGGAGACCTGTCTTACTCAGACACACACAGTGT	1440		
QY	1712	ACAGAGCAGACACACAGATCTTCCAGATCAGCAGGAGAGACCCCGGAGCCTCTGCTCTC	1771		
DB	1441	ACAGAGCAGACACACAGATCTTCCAGATCAGCAGGAGAGACCCCGGAGCCTCTGCTCTC	1500		
QY	1772	CTACACTGGCATGCTGATGAGATCGTGAATGCCCAATTTGGCTTCTTCCACATCTGGTT	1831		
DB	1501	CTACACTGGCATGCTGATGAGATCGTGAATGCCCAATTTGGCTTCTTCCACATCTGGTT	1560		
QY	1832	GCATCTGCTATGATGGCTCGCTGCATCTCCCTCAGTCCCAAAATTTCTAGTAGCCAGTGT	1891		
DB	1561	GCATCTGCTATGATGGCTCGCTGCATCTCCCTCAGTCCCAAAATTTCTAGTAGCCAGTGT	1619		
QY	1892	TCCTGCGAGGCTGTCTATGCTGCTGCTGCTGCCAAGGAGACACTCTTCAGAGCAATTTT	1951		
DB	1620	TCCTGCGAGGCTGTCTATGCTGCTGCTGCTGCCAAGGAGACACTCTTCAGAGCAATTTT	1678		
QY	1952	TGGGTAAAGGAACACTTACAAAGAGGCAATTTGATCTTGTGCTGAGGCTCAGAGCCCTTTT	2011		
DB	1679	TGGGTAAAGGAACACTTACAAAGAGGCAATTTGATCTTGTGCTGAGGCTCAGAGCCCTTTT	1738		
QY	2012	GATAGGCTCTGATGCTCATTAAGACATTTCAAGCCAGATGCTCCAACTGCAAAATAT	2071		
DB	1739	GATAGGCTCTGATGCTCATTAAGACATTTCAAGCCAGATGCTCCAACTGCAAAATAT	1797		
QY	2072	ACCAACCTCTCTGAATATATTTTGTATTTATTTATTTCTTTTCTTTTCTTAAAGAA	2131		
DB	1798	ACCAACCTCTCTGAATATATTTTGTATTTATTTATTTCTTTTCTTTTCTTAAAGAA	1857		
QY	2132	TTGGCTCTGAATAGATGCAATTTTCCATCTGACCTGAGTGCATATCAATTTAGCCATC	2191		

Db 1858 -TGGCTCTGAATAGATGCACATTTTCCAT-TGAACGTGATGCATTTCAATTAGCAATC 1915
QY 2192 CAGTAATTTATTTATTAATATATACATAATATATGTTTCTCAGCATAGGAGCTATGATT 2251
Db 1916 CAGTAATTTATTTATTAATATATACATAATATATGTTTCTCAGCATAGGAGCTATGATT 1975
QY 2252 CATTAAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTTATTTTCATTACA 2311
Db 1976 CATTAAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTTATTTTCATTACA 2035
QY 2312 CAAACTTAATTTGTTCTTGTAAATTAAGTTCAAGTGGATCTTCGAGTGGGATTTCTTGGA 2371
Db 2036 CAAACTTAATTTGTTCTTGTAAATTAAGTAC-ASTGGAUCTTCGAGTGGGATTTCTTGGA 2094
QY 2372 AATTATCTTGCACTTGAATGTCATGATTACATATGAATCGCTTTGACATATCTTTAG 2431
Db 2095 AATTATCTTGCACTTGAATGTCATGATTACATATGAATCGCTTTGACATATCTTTAG 2154
QY 2432 ACAGAAAAAGTAGCTAGTGGAGGGGAAATATATAGAGCTTGTGACTTTAGGAGTAG 2491
Db 2155 ACAGAAAAAGTAGCTAGTGGAGGGGAAATATATAGAGC-TGTGTACTTTAGGAGTAG 2213
QY 2492 CT 2493
Db 2214 GT 2215
RESULT 11
ADC39137
ID ADC39137 standard; cDNA; 2331 BP.
AC ADC39137;
XX
DT 18-DEC-2003 (first entry)
XX
DE Novel human NOVX polypeptide coding sequence SEQ ID NO: 79.
XX
KW ds; gene; antidiabetic; cytostatic; immunomodulator; anorectic;
KW antilipemic; nootropic; neuroprotective; immunostimulant;
KW antiparkinsonian; anti-HIV; antiasthmatic; antiinflammatory; hypotensive;
KW antiarteriosclerotic; hemostatic; osteopathic; gene therapy.; NOVX;
KW diabetes; obesity; cancer; lymphoma; uterus cancer; prostate cancer;
KW dyslipidemia; anorexia; wasting disorder; Alzheimer's disease;
KW Parkinson's disorder; cachexia; cardiomyopathy; AIDS; asthma;
KW Crohn's disease; multiple sclerosis; hypertension; atherosclerosis;
KW hemophilia; graft-versus-host disease;
KW Albricht hereditary osteodystrophy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 16..769
FT /*tag= a
PN
PN WO2003010327-A2.
XX
PD 06-FEB-2003.
XX
PF 02-MAY-2002; 2002WO-US014199.
XX
PR 02-MAY-2001; 2001US-0288063P.
PR 03-MAY-2001; 2001US-0288395P.
PR 07-MAY-2001; 2001US-0289087P.
PR 09-MAY-2001; 2001US-0289817P.
PR 09-MAY-2001; 2001US-0289818P.
PR 11-MAY-2001; 2001US-0290194P.
PR 14-MAY-2001; 2001US-0290753P.
PR 15-MAY-2001; 2001US-0291181P.
PR 16-MAY-2001; 2001US-0291243P.
PR 18-MAY-2001; 2001US-0292001P.
PR 21-MAY-2001; 2001US-0292374P.
PR 22-MAY-2001; 2001US-0292587P.

PR 23-MAY-2001; 2001US-0293107P.
PR 25-MAY-2001; 2001US-0293747P.
PR 29-MAY-2001; 2001US-0294109P.
PR 29-MAY-2001; 2001US-0294110P.
PR 30-MAY-2001; 2001US-029434P.
PR 31-MAY-2001; 2001US-0294827P.
PR 12-JUL-2001; 2001US-0304879P.
PR 31-JUL-2001; 2001US-0308901P.
PR 14-AUG-2001; 2001US-0312270P.
PR 17-AUG-2001; 2001US-0313416P.
PR 10-SEP-2001; 2001US-0318463P.
PR 27-SEP-2001; 2001US-0325683P.
PR 18-OCT-2001; 2001US-0330292P.
PR 28-NOV-2001; 2001US-0333873P.
PR 03-DEC-2001; 2001US-0336909P.
PR 03-DEC-2001; 2001US-0337552P.
PR 21-FEB-2002; 2002US-0359245P.
PR 01-MAY-2002; 2002US-00136826.
XX
XX (CURA-) CURAGEN CORP.
PA Miller CE, Kekuda R, Malyankar UM, Li L, Pena CEA, Spytek KA;
PI Gorman L, Guo X, Fernandes ER, Smithson G, Stone DJ, Zerkhusen BD;
PI Patturajan M, Anderson DW, Mezes PS, Peyman JA, Macdougall JR;
PI Padigaru M, Rastelli L, Shenoy SG, Gerlach VL, Shinkets RA, Zhong M;
PI Edinger SR, Ellerman K;
XX
XX WPI; 2003-239445/23.
DR P-PSDB; ADC39138.
XX
PT New NOVX polypeptides and polynucleotides, useful in gene therapy, a
PT particularly for treating or preventing a syndrome associated with a
PT human disease e.g. diabetes, obesity, cancer, Alzheimer's disease,
PT hypertension or hemophilia.
XX
PS Claim 20; SEQ ID NO 79; 748pp; English.
XX
CC The invention relates to new isolated NOVX polypeptides, the genes
CC encoding them or sequences having at least 95% identity to the amino acid
CC or nucleotide sequences. The NOVX polypeptide is useful as a therapeutic,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The NOVX polypeptide is particularly useful for
CC treating, preventing or alleviating pathology associated with NOVX
CC polypeptide in a mammal, e.g. a human. The NOVX nucleic acid and
CC polypeptide are especially useful for treating or preventing e.g.
CC diabetes, obesity, cancers (e.g. lymphoma, uterus cancer or prostate
CC cancer), dyslipidemias, anorexia, wasting disorders, Alzheimer's disease,
CC Parkinson's disorder, cachexia, cardiomyopathy, AIDS, asthma, Crohn's
CC disease, multiple sclerosis, hypertension, atherosclerosis, hemophilia,
CC graft-versus-host disease or Albricht hereditary osteodystrophy. The DNA
CC encoding the protein is useful in gene therapy for treating the above
CC conditions. These are also useful in developing powerful assay system for
CC functional analysis of various human disorders, as well as in diagnostic
CC applications. This sequence represents one of the NOVX genes of the
CC invention.
XX
SQ Sequence 2331 BP; 646 A; 502 C; 497 G; 686 T; 0 U; 0 Other;
Query Match 66.3%; Score 1704.8; DB 10; Length 2331;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1795; Conservative 0; Mismatches 22; Indels 7; Gaps 7;
QY 670 CCCGACCGGGAAGCGCCCGCGGCGTTTGTGAGAGGTGTACCTCTTTATTCCTGAAGC 729
Db 60 CTTGGCTCACTTCGCCCTAGCGCGTTTGTGAGAGGTGTACCTCTTTATTCCTGAAGC 119
QY 730 CCGGCACTGTCTGACCTGTGTCGACCTGCGCCAAAACAGATCCAGAAAGAAATTCGA 789
Db 120 CCGGCACTGTCTGACCTGTGTCGACCTGCGCCAAAACAGATCCAGAAAGAAATTCGA 179
QY 790 GGTGTTGCCATCGAAGGACTGGATCCACGGGTAAACACCGGTGACCCAGTCAGTCGC 849

Db	780	GCACGTCGCCACGCTAACTAGATTAGATTGTTGTTGAAACATCTACATCCACCACTTTGT	839
QY	1450	TATGAGTGTTCCTCCAAATTCCTGTTCTACAAGCATGTTGTGGCAGAAACCTGGAGACC	1509
Db	840	TATGAGTGTTCCTCCAAATTCCTGTTCTACAAGCATGTTGTGGCAGAAACCTGGAGACC	899
QY	1510	AGGCATCTTAATTTACTTCAGCCATCGTACCTCTTCGACATGATGACCCGTCATCAC	1569
Db	900	AGGCATCTTAATTTACTTCAGCCATCGTACCTCTTCGACATGATGACCCGTCATCAC	959
QY	1570	AAAGTCCCTCTCATCATCTCCAGTCGAGAGCCAGCGATTGCTTTCTTCCTGGCATAGT	1629
Db	960	AAAGTCCCTCTCATCATCTCCAGTCGAGAGCCAGCGATTGCTTTCTTCCTGGCATAGT	1019
QY	1630	AAACATTTTCTTGGAAATATGTTTCACTTAATCACTACCAATATCTCGAAGACCTGTC	1689
Db	1020	AAACATTTTCTTGGAAATATGTTTCACTTAATCACTACCAATATCTCGAAGACCTGTC	1079
QY	1690	TTACTCAGACAGACACAGGTGTACAGAACAGACAGACAGATCTTCAGATCAGCAGGA	1749
Db	1080	TTACTCAGACAGACACAGGTGTACAGAACAGACAGACAGATCTTCAGATCAGCAGGA	1139
QY	1750	GACCCCGGAGCCTCTCTCTCTACACTGGCATGCTGATGAGATCGTGACATGCCACACA	1809
Db	1140	GACCCCGGAGCCTCTCTCTCTACACTGGCATGCTGATGAGATCGTGACATGCCACACA	1199
QY	1810	TTGGCTTCTTCCACATCTGGTTCGACTCGTCACTGATGGCTCGCTGCAATCCCTCAGTC	1869
Db	1200	TTGGCTTCTTCCACATCTGGTTCGACTCGTCACTGATGGCTCGCTGCAATCCCTCAGTC	1259
QY	1870	CCAAATCTAGTACGACGATGTTCTGACAGAGCTGTCTGATGCTGCTGCTGCTGCTGCT	1929
Db	1260	CCAAATCTAGTACGACGATGTTCTGACAGAGCTGTCTGATGCTGCTGCTGCTGCTGCT	1317
QY	1930	GACACTCTTCAGAGCCTTTTGGTAAAGGAAACACTTACAAAGAGGCTTGTATCTGT	1989
Db	1318	GACACTCTTCAGAGCCTTTTGGTAAAGGAAACACTTACAAAGAGGCTTGTATCTGT	1377
QY	1990	GTCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTGATGCTGATGCTGATGCTGAT	2049
Db	1378	GTCTGAGGCTCAGAGCCCTTTTGTAGGCTTCTGATGCTGATGCTGATGCTGATGCTGAT	1436
QY	2050	AGATGCTCAACTGCAAAATATACCAACCTTCTGATATATATTTTGTATTTATATTT	2109
Db	1437	AGATGCTCAACTGCAAAATATACCAACCTTCTGATATATATTTTGTATTTATATTT	1496
QY	2110	TCCTTTCTTTTCTTAAAGAAATGGCTCTGAATAGAGTACACATTTTCCATCTGAATG	2169
Db	1497	TCCTTTCTTTTCTTAAAGTAA-TGGCTCTGAATAGAGTACACATTTTCCAT-TGAATG	1554
QY	2170	GATGCATATCATTTAGCCATPCAGTAATTTATTTATTTATTTATTTATTTATTTATTT	2229
Db	1555	GATGCATTTTATTTAGCCATPCAGTAATTTATTTATTTATTTATTTATTTATTTATTT	1614
QY	2230	CCTCAGATAGGACTATGATTTCAATTAATTAAGTGGAGTCAAAACCTTAAATGCAATG	2289
Db	1615	CCTCAGATAGGACTATGATTTCAATTAATTAAGTGGAGTCAAAACCTTAAATGCAATG	1674
QY	2290	TTTGTGTGTATTTTCAATACACAAACCTTAAATTTGTCTTTGTTAAATTAAGTTCAAGT	2349
Db	1675	TTTGTGTGTATTTTCAATACACAAACCTTAAATTTGTCTTTGTTAAATTAAGTTCAAGT	1733
QY	2350	CTTGAGTGGGATTTCTGGTAAATTTATCTTGACCTTGAATGTCTCATGATACATATGA	2409
Db	1734	CTTGAGTGGGATTTCTGGTAAATTTATCTTGACCTTGAATGTCTCATGATACATATGA	1793
QY	2410	AATGCTTTTGACATATCTTTTACAGAGAAAGTACCTGAGTGGGGGAAATTTATAGAG	2469
Db	1794	AATGCTTTTGACATATCTTTTACAGAGAAAGTACCTGAGTGGGGGAAATTTATAGAG	1853
QY	2470	CTTGTGTGACTTTAGGAGTAGCT	2493
Db	1854	C-TGTGTGACTTTAGGAGTAGGT	1876

RESULT 13

ADF76970

ID ADF76970 standard; cDNA; 2566 BP.

XX ADF76970;

AC ADF76970;

XX 26-FEB-2004 (first entry)

DT 26-FEB-2004 (first entry)

XX Novel human secreted and transmembrane protein cDNA SeqID 645.

DE human; PRO; membrane bound protein; membrane bound receptor;

XX cell proliferation; cell migration; cell differentiation;

XX mitogenic factor; survival factor; cytotoxic factor;

XX differentiation factor; neuropeptide; hormone; cell receptor;

XX receptor-ligand interaction; cytoskeletal; chondrocyte; tumour; ss; gene.

XX Homo sapiens.

OS Homo sapiens.

XX WO2003072035-A2.

PN 04-SEP-2003.

XX 21-FEB-2003; 2003WO-US005241.

XX 22-FEB-2002; 2002US-0359461P.

PR (GETH) GENENTECH INC.

XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;

PI Williams PM, Wood WI, Wu TD;

XX WPI; 2003-721702/68.

DR P-PSDB; ADF76971.

XX New PRO polypeptides, useful for diagnosing and treating an immune

XX related disorder, e.g. systemic lupus erythematosus, rheumatoid

XX arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or

XX diabetes mellitus.

XX Claim 2; SEQ ID NO 645; 918bp; English.

XX This invention relates to novel nucleic acids encoding human PRO secreted

XX and transmembrane proteins. Extracellular proteins play important roles

XX in the formation, differentiation and maintenance of multicellular

XX organisms. The fate of many individual cells (for example proliferation,

XX migration or differentiation) is typically governed by information

XX received from other cells and the immediate environment. The information

XX is often transmitted by secreted polypeptides (for example mitogenic

XX factors, survival factors, cytotoxic factors, differentiation factors,

XX neuropeptides and hormones) which are received and interpreted by diverse

XX cell receptors or membrane bound proteins. These membrane bound proteins

XX and receptors may be of use as pharmaceutical and diagnostic agents, such

XX as in the blocking of receptor-ligand interactions. The current invention

XX provides the amino acid sequences of novel human membrane bound receptors

XX and proteins, along with the cDNA sequences encoding them. The novel

XX proteins of the invention may have cytostatic activities through the

XX stimulation of chondrocytes. The nucleic acids of the invention may be

XX useful for the manufacture of a medicament for diagnosing or treating a

XX tumour in a mammal. In addition, they may be useful for measuring or

XX detecting the expression of a tumour associated gene. The present

XX sequence is a cDNA sequence which encodes a human PRO protein of the

XX invention.

XX Sequence 2566 BP; 720 A; 520 C; 565 G; 761 T; 0 U; 0 Other;

XX SQ

Query Match 65.9%; Score 1694.2; DB 10; Length 2566;

Best Local Similarity 98.9%; Pred. NO. 0;

Matches 1779; Conservative 0; Mismatches 13; Indels 7; Gaps 7;

QY 695 GTTTGGAGGAGTGTACCTCTTTATTTCTGAAGCCCGGAGTCTTGACCTGTCGAC 754

Db	187	GCTGTGCACAGTGA	CCTCCTTTATTTCTGGAAGCCCGGGCAGTGCCTTTGACCTGGTGCAC	246
Qy	755	CAGTGC	CCAAAACACAGATCCAGAAAGGAAAGTTTCCAGGTGTGTGCCATCGAAGGACTGGAT	814
Db	247	CAGTGC	CCANAACAGATCCAGAAGGAAAGTTTCCAGGTGTGTGCCATCGAAGGACTGGAT	306
Qy	815	GCCACGGGTAAAC	ACACCGGTGACCCAGTCACTGGCAGATTCACTTAAGGCTGTCTCTTTA	874
Db	307	GCCACGGGTAAAC	ACACCGGTGACCCAGTCACTGGCAGATTCACTTAAGGCTGTCTCTTTA	366
Qy	875	AAGTCACCA	CCCTCTTGTCATTTGGCCAGTGGAGGAGATCTTTGATGATGAACCAACTATC	934
Db	367	AAGTCACCA	CCCTCTTGTCATTTGGCCAGTGGAGGAGATCTTTGATGATGAACCAACTATC	426
Qy	935	ATTAGAAGAG	CTTTTACTCTTTGGGCAATTAATTTGGGCTCCGGAATAGCTAAAGAA	994
Db	427	ATTAGAAGAG	CTTTTACTCTTTGGGCAATTAATTTGGGCTCCGGAATAGCTAAAGAA	486
Qy	995	TCTGCCAAAT	CTCTGTGATTTGTAGACAGGTACTGGCACACGACCGGCCA	1054
Db	487	TCTGCCAAAT	CTCTGTGATTTGTAGACAGGTACTGGCACACGACCGGCCA	546
Qy	1055	GCCACTCAG	GTGAGTGGGGTCTCCAGCACTGCGCCCCCAGGCCCATCACCCTGTGTACGAG	1114
Db	547	GCCACTCAG	GTGAGTGGGGTCTCCAGCACTGCGCCCCCAGGCCCATCACCCTGTGTACGAG	606
Qy	1115	TGGCCAGAG	CACTGTCTCAAACTGACCTTATCTCTGTGCTCTCACTGTGAGTCTCTGAGGAG	1174
Db	607	TGGCCAGAG	CACTGTCTCTCAAACTGACCTTATCTCTGTGCTCTCACTGTGAGTCTCTGAGGAG	666
Qy	1175	AGTTTGACAG	GGCTGCAGGGCCGGGGCATGGAGAAGACCGAGGAAGACAGAACTTTGAG	1234
Db	667	AGTTTGACAG	GGCTGCAGGGCCGGGGCATGGAGAAGACCGAGGAAGACAGAACTTTGAG	726
Qy	1235	GCCAAACAG	TGTTTCGTCTCAAAAGGTAGAAATGTCCTACAGCGGATGGAGAATCCTGGC	1294
Db	727	GCCAAACAG	TGTTTCGTCTCAAAAGGTAGAAATGTCCTACAGCGGATGGAGAATCCTGGC	786
Qy	1295	TGCCATGTG	TTGATGCCAGCCCTCCAGAGAAAAGTCTCTGCAGACAGATTTAAGGCTA	1354
Db	787	TGCCATGTG	TTGATGCCAGCCCTCCAGAGAAAAGTCTCTGCAGACAGATTTAAGGCTA	846
Qy	1355	ATCCAGAA	TAGTTTATGTAAACCGTAGTTACTCTGGCCAGGTGCCACGCTCTAACTAGATT	1414
Db	847	ATCCAGAA	TAGTTTATGTAAACCGTAGTTACTCTGGCCAGGTGCCACGCTCTAACTAGATT	906
Qy	1415	AGATGTTG	TTTGAACATCTACATCCACCATTTTGTTATGCAGTGTCCCAAAATTTCTGTT	1474
Db	907	AGATGTTG	TTTGAACATCTACATCCACCATTTTGTTATGCAGTGTCCCAAAATTTCTGTT	966
Qy	1475	CTACAAG	CAATGTTGTGGCAGAAAACCTGGAGACCGAGCATCTTAATTTTACTTCAGCCA	1534
Db	967	CTACAAG	CAATGTTGTGGCAGAAAACCTGGAGACCGAGCATCTTAATTTTACTTCAGCCA	1026
Qy	1535	TCGTAC	CCCTCTTCTGATGTGACCCGTCATCAAAAGTCCCTCTCATCATGTTCCAG	1594
Db	1027	TCGTAC	CCCTCTTCTGATGTGACCCGTCATCAAAAGTCCCTCTCATCATGTTCCAG	1086
Qy	1595	TGAGAG	CCAGCGATTCTTCTTCTCGGCATAGTAACATTTTCTTTGGAACATATGTTT	1654
Db	1087	TGAGAG	CCAGCGATTCTTCTTCTCGGCATAGTAACATTTTCTTTGGAACATATGTTT	1146
Qy	1655	CACTTAAT	CACTACCAATATCTGGAAGACCTGTCTTTACTTCAGACAGCACCAAGTGTACA	1714
Db	1147	CACTTAAT	CACTACCAATATCTGGAAGACCTGTCTTTACTTCAGACAGCACCAAGTGTACA	1206
Qy	1715	GAAGCAG	GACAGATCTTCCAGATCAGCAGGGAGACCCCGGAGCCTCTGCTTCTCCTA	1774
Db	1207	GAAGCAG	GACAGATCTTCCAGATCAGCAGGGAGACCCCGGAGCCTCTGCTTCTCCTA	1266
Qy	1775	CACTTG	CACTGTATGATGATCGTGACATGCCCATTTGGGCTTCTTCCATCTCGGTTGCA	1834
Db	1267	CACTTG	CACTGTATGATGATCGTGACATGCCCATTTGGGCTTCTTCCATCTCGGTTGCA	1326

Qy	1835	CTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATTCCTAGTAGCCAAAGTGTTCC	1894
Db	1327	CTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATTCAG - AGCCAAAGTGTTCC	1385
Qy	1895	TGCAGAGGCTGCTCTATGTGTGCTCGCTGCCCAAGGGACACTCCTGCAGAGCCATTTTTGG	1954
Db	1386	TGCAGAGGCTGCTATGTGTGCTCGCTGCCCTGCCA - GGAACATCTCTGCAGAGCCATTTTTGG	1444
Qy	1955	GTAAGGAACACTTACAAAGGAAGGCATTGATCTTGTGTCTGAGGCTCAGAGCCCTTTTGAT	2014
Db	1445	GTAAGGAACACTTACAAAGGAAGGCATTGATCTTGTGTCTGAGGCTCAGAGCCCTTTTGAT	1504
Qy	2015	AGGCTCTTGATGTCATTCATAAGACACATCAAGCCRAAGATGCTCCCACTGCAATATACC	2074
Db	1505	AGGCTCTTGA - GTCATATATAAGACACATCAAGCCRAAGATGCTCCCACTGCAATATACC	1563
Qy	2075	AACCTTCTCGAAATTAATTTTGGTATTTATATTTCTTTTCTTTTCTTTTCTTAAAGAAATG	2134
Db	1564	AACCTTCTCGAAATTAATTTTGGTATTTATTTCTTTTCTTTTCTTTTCTTAAAGTA - TG	1622
Qy	2135	GCTCTGAATAGATGCAATTTTCCATCTGAACTGGATGCATATCATTTAGCCAAATCCAG	2194
Db	1623	GCTCTGAATAGATGCAATTTTCCAT - TGAACCTGGATGCATTTCAATTTAGCCAAATCCAG	1681
Qy	2195	TAATTTATTTATATTAATCTATACATAAATATGTTTCTCAGCATAGAGAGTATGATTCAT	2254
Db	1682	TAATTTATTTATATTAATCTATACATAAATATGTTTCTCAGCATAGAGAGTATGATTCAT	1741
Qy	2255	TAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTATTTTCAATACACAA	2314
Db	1742	TAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTATTTTCAATACACAA	1801
Qy	2315	ACTTAATTTGCTTGTTTAAATAGTTCACAGTGCATCTTGAGTGGGATTTTCTTGGTAAAT	2374
Db	1802	ACTTAATTTGCTTGTTTAAATAGTAC - AGTGATCTTGAGTGGGATTTTCTTGGTAAAT	1860
Qy	2375	TATCTTGCACTTGAATCTGTCATGATTACATATGAATCGCTTTGCATATCTTTAGACA	2434
Db	1861	TATCTTGCACTTGAATGCTCATGATTACATATGAATCGCTTTGCATATCTTTAGACA	1920
Qy	2435	GAATAAGTAGCTGAGTGGGGGAAATTAAGAGCTTGTGTGACTTTAGGAGTAGCT	2493
Db	1921	GAATAAGTAGCTGAGTGGGGGAAATTAAGAGC - TGTGTCACTTTAGGAGTAGGT	1978

RESULT 14

RESOL1 14
ABV26811
ID ABV26811 standard; cDNA; 2199 BP.

AC ABV26811:

XX
DT 16-SEP-2002 (first entry)XX
DE Human prostate expression marker cDNA 26802.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

xxviii Parameter, marker, gene, base

US Homo sapiens.
yy

PN WO200160860-A2.

PD 23-AUG-2001.

20-FEB-2001: 2001WO-US005171

XX
PB 17-FEB-2000, 2000US-0193318D

PR 16-MAR-2000; 2000US-0189862P.

PR 09-JUN-2000; 2000US-0211314P.

I8-D007-2000; 2000US-0213001E
 I3-DEC-2000; 2000US-0255281P
 PR

Db 1851 TTATCTTGCACTTGAAATGCTCTCTGATTACATATGAATCGC-TTGACATATCTTTGGAC 1309
QY 2434 AGAAAAAGTAGCTGAGTGAGGGGAAATTAAGAGCTTGTGTGACATTTAGGGAGTAGCT 2493
Db 1910 AGAAAAAGTAGTGAGTGAGGGGAAATTAAGAGCTTGTGTGACATTTAGGGAGTAGGT 1969
RESULT 15
ABV20963
ID ABV20963 standard; cDNA; 2199 BP.
XX AC ABV20963;
XX AC
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 20954.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WC200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 3446; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 2199 BP; 597 A; 488 C; 516 G; 592 T; 0 U; 6 Other;
Query Match 61.4%; Score 1579.6; DB 5; Length 2199;
Best Local Similarity 97.1%; Pred. No. 2.5e-306;
Matches 1748; Conservative 0; Mismatches 34; Indels 18; Gaps 13;
QY 695 GTTTGGAGGAGTGACCTCTTTATTCTGAAGCCCGGCGAGTGCTTGACCTGTCGAC 754
Db 187 GCTTGTGACATGTACCTCTTTATTCTGAAGCCCGGCGAGTGCTTGACCTGTCGAC 246
QY 755 CAGTCCCAAAACAGATCCAGAAAGAAAGTTCCAGGTTGTTGCCATCGAAGGACTGGAT 814
Db 247 CAGTCCCAAAACAGATCCAGAAAGAAAGTTCCAGGTTGTTGCCATCGAAGGACTGGAT 306

QY 815 GCCACGGTAAAAACACCGGTGACCCAGTCAGTGCAGATTCACCTTAAGGCTGTCTCTTA 874
Db 307 GCCACGGTAAAAACACCGGTGACCCAGTCAGTGCAGATTCACCTTAAGGCTGTCTCTTA 366
QY 875 AAGTCACCAACCTCTTTCGATTTGGCCAGTGGAGGAAGATCTTTGATGATGAACCACTATC 934
Db 367 AAGTCACCAACCTCTTTCGATTTGGCCAGTGGAGGAAGATCTTTGATGATGAACCACTATC 426
QY 935 ATTAGAAGAGCTTTTACTCTTTGGGCAATTAATTTGGCCCTCCGAAATAGCTAAAGAA 994
Db 427 ATTAGAAGAGCTTTTACTCTTTGGGCAATTAATTTGGCCCTCCGAAATAGCTAAAGAA 486
QY 995 TCTGCCAAATCTCTGTGATTGTAGACAGTACTGGACACAGCGGCCACCTATGCCATA 1054
Db 487 TCTGCCAAATCTCTGTGATTGTAGACAGTACTGGACACAGCGGCCACCTATGCCATA 546
QY 1055 GCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCCCAGCCCATCACCTGTGTACCAG 1114
Db 547 GCCACTGAGGTGAGTGGGGTCTCCAGCACCTGCCCCCAGCCCATCACCTGTGTACCAG 606
QY 1115 TGGCCAGAGGACCTGCTCAAAACCTGACCTTATCTCTGCTGCTACTGTGAGTCTCTGAGGAG 1174
Db 607 TGGCCAGAGGACCTGCTCAAAACCTGACCTTATCTCTGCTGCTACTGTGAGTCTCTGAGGAG 666
QY 1175 AGTTTCAGAGGCTGCAGGGCCGGGCATGGAGAAGACCAGGGAAGAAAGCAGAACTTGAG 1234
Db 667 AGTTTCAGAGGCTGCAGGGCCGGGCATGGAGAAGACCAGGGAAGAAAGCAGAACTTGAG 726
QY 1235 GCCAACAGTGTGTTTCGTCAAAAGGTAGAAATGCTCTACACGGGATGAGATCTCTGCC 1294
Db 727 GCCAACAGTGTGTTTCGTCAAAAGGTAGAAATGCTCTACACGGGATGAGATCTCTGCC 786
QY 1295 TGCATGTGTTGATGCCAGCCCTCCAGAGAAAAGCTCTGCAGACAGTATTAAAGCCTA 1354
Db 787 TGCATGTGTTGATGCCAGCCCTCCAGAGAAAAGCTCTGCAGACAGTATTAAAGCCTA 846
QY 1355 ATCCAGATAGTATTAGTGAAACCGTAGTTACTCTGCCCAGGTGCCACGTCCTAAGATT 1414
Db 847 ATCCAGATAGTATTAGTGAAACCGTAGTTACTCTGCCCAGGTGCCACGTCCTAAGATT 906
QY 1415 AGATGTGTTTCAAAACATCTACATCCACATTTGTTTATCGAGTGTCCCAAAATTTCTGTT 1474
Db 907 AGATGTGTTTCAAAACATCTACATCCACATTTGTTTATCGAGTGTCCCAAAATTTCTGTT 966
QY 1475 CTACAAGCATGTTGTGTGGCAGAAAACCTGGAGACAGGATCTTAATTTTACTTTCAGCCA 1534
Db 967 CTACAAGCATGTTGTGTGGCAGAAAACCTGGAGACAGGATCTTAATTTTACTTTCAGCCA 1026
QY 1535 TCGTACCTCTTCTGACTGATGAGCCCGTCATCAAAAGGTCCCTCTCATCATGTTCCAG 1594
Db 1027 TCGTACCTCTTCTGACTGATGAGCCCGTCATCAAAAGGTCCCTCTCATCATGTTCCAG 1086
QY 1595 TGAGAGGCGCAGGATGCTTTCTTCTGCGATAGTAAACATTTCTTGGAAATATGTTT 1654
Db 1087 TGAGAGGCGCAGGATGCTTTCTTCTGCGATAGTAAACATTTCTTGGAAATATGTTT 1146
QY 1655 CACTTAATCACTACCAAAATATCTGGAAGACCTGTCTTACTCAGACAGACAGGTTGACA 1714
Db 1147 CACTTAATCACTACCAAAATATCTGGAAGACCTGTCTTACTCAGACAGACAGGTTGACA 1206
QY 1715 GAAGCAGCAGCAAGATCTTCCAGATCAGAGGGAGACCCCGAGGACCTTCTGCTCTCTTA 1774
Db 1207 GAAGCAGCAGCAAGATCTTCCAGATCAGAGGGAGACCCCGAGGACCTTCTGCTCTCTTA 1266
QY 1775 CACTGGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTTCCACATCTGTTGCA 1834
Db 1267 CACTGGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTTCCACATCTGTTGCA 1326
QY 1835 CTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATTTCTAGTACCAAGGTTTCC 1894
Db 1327 CTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAAATTTCTAG-AGCCAAAGTGTTC 1385

QY	1895	TGCAGAGCTGTCTATGTGTCTGTGCTGCCAAGGACACTCTCTGCAGAGCCATTTTGG	1954
Db	1386	TGCAGAGGCTGTCTATGTGTCTGTGCTGCCAA-GGACACTCTGCAGAGCCATTTTGG	1444
QY	1955	GTAAGGACACTTACAAAGGCAATGTCTGTGTCTGAGGCTCAGAGCCCTTTTGAT	2014
Db	1445	GTAAGGACACTTACAAAGGCAATGTCTGTGTCTGAGGCTCAGAGCCCTTTTGAT	1504
QY	2015	AGGCTTCTCAUGTCAATTCATAAGACATTCAGCCAAAGTCCAACTGCAAAATATACC	2074
Db	1505	AGGCTTCTGA-GTCATATATAAGACATTCAGCCAAAGTCCAACTGCAAAATATACC	1563
QY	2075	AACCTTCTCTGAATATATTTTGGCTTATTTATATTTCTTTCTTTTCTTAAGAATTG	2134
Db	1564	AA-CTTCTCTGAATTATA-TTTGCTTATTATATTTCTTTCTTTTCTTAAGTA--TG	1619
QY	2135	GCTCTGAATAGATGACATTTTCCATCTGAAGTGGATGCATATCATTTAGCCAAATCCAG	2194
Db	1620	GCTCTGAATAGATGACATTTTCCAT-TGAAGTGGATGCATTTCAATTTAGCCAAATCCAG	1678
QY	2195	TAATTTTATTATTAATCTATACATAATATGTTTCTCAGCATAGGAGCTATGATTCAT	2254
Db	1679	TAATTTTATTATTAATCTATACATAATATGTTTCTCAGCATAGGAGCTATGATTCAT	1738
QY	2255	TAATTTAAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTATTTTTCATTACACAA	2314
Db	1739	TAATTTAAAGTGGAGTCAAAA-GCTAAATGCAATGTTTGTGTATTTTTCATTACACAA	1794
QY	2315	ACTTAATTTCTTGTGTTTAAATAAGTTCAGTGGATCTTGGAGTGGGATTTCTT-GGTAAA	2373
Db	1795	A--TTAATTTCTTGTGTTTAAATAAGTTCAGTGGATCTTGGAGTGGGATTTCTTGGGTAAA	1850
QY	2374	TTATCTTGCACCTTGAATGCTCTCATGATTACATATGAATCGCTTTGACATATCTTTAGAC	2433
Db	1851	TTATCTTGCACCTTGAATGCTCTCTGATTACATATGAATCGC-TTGACATATCTTTGGAC	1909
QY	2434	AGAAAAAGTAGCTGAGTGGGGGAAATTTATAGAGCTTGTGACCTTTAGGGAGTAGCT	2493
Db	1910	AGAAAAAGTAGCTGAGTGGGGGAAATTTATAGAGGCTGTGTGACCTTTAGGGAGTAGGT	1969

Search completed: October 22, 2004, 15:07:42
Job time : 1235 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 08:04:13 ; Search time 10840 Seconds

(without alignments)
11216.035 Million cell updates/sec

Title: US-10-681-223-1

Perfect score: 2571

Sequence: 1 cggggccaggcggtcg.....aggaggggcttttgcaccc 2571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2571	100.0	2571	6	AR258675 Sequence
2	2571	100.0	2571	6	AR438694 Sequence
3	1579.6	61.4	2199	6	CQ489087 Sequence
4	1579.6	61.4	2199	6	CQ490198 Sequence
5	1579.6	61.4	2199	6	CQ494933 Sequence
6	1579.6	61.4	2199	6	CQ496057 Sequence
7	1459.6	56.8	1535	6	AX814311 Sequence
8	1336.6	52.0	1343	6	AX814308 Sequence
9	1193.8	46.4	1720	9	BC016969 Homo sapi
10	1141.4	44.4	118808	9	AC017076 Homo sapi
11	1130.4	44.0	20966	6	AR258676 Sequence
12	1130.4	44.0	20966	6	AR438695 Sequence
13	1065.2	41.4	1215	6	CQ733692 Sequence
14	889.2	34.6	1609	9	AK127983 Homo sapi
15	873	34.0	3028	10	BC057565 Mus muscu
16	797.2	31.0	1188	6	CQ777457 Sequence
17	690.2	26.8	757	6	AX814328 Sequence
18	674.8	26.2	759	6	AX814327 Sequence
19	646.6	25.1	833	6	AX814312 Sequence

20	621.8	24.2	844	6	AX814319 Sequence
21	600.4	23.4	972	6	AX814313 Sequence
22	569	22.1	728	6	AX814321 Sequence
23	562	21.9	700	6	AX814315 Sequence
24	562	21.9	700	6	AX814322 Sequence
25	534	20.8	702	6	AX814323 Sequence
26	533.2	20.7	795	6	AX814316 Sequence
27	524.2	20.4	669	6	AX814329 Sequence
28	506	19.7	1008	6	AX814318 Sequence
29	493.8	19.2	928	6	AX814326 Sequence
30	482.6	18.8	556	6	AX814330 Sequence
31	480.4	18.7	674	6	AX814317 Sequence
32	463.6	18.0	3292	6	AX306082 Sequence
33	463.6	18.0	3292	10	MUSTHYKIN
34	456.6	17.8	633	6	AX814314 Sequence
35	453.4	17.6	531	6	CQ776571 Sequence
36	411.4	16.0	454	6	CQ501652 Sequence
37	411.4	16.0	454	6	CQ508973 Sequence
38	411.4	16.0	454	6	CQ510575 Sequence
39	411.2	16.0	418	6	CQ480481 Sequence
40	399	15.5	417	6	CQ478818 Sequence
41	390.8	15.2	454	6	CQ471313 Sequence
42	382.8	14.9	432	6	CQ500010 Sequence
43	360.6	14.0	443	11	G23005
44	360	14.0	226460	2	AC117652
45	360	14.0	309317	2	AC132190

ALIGNMENTS

RESULT 1
AR258675
LOCUS AR258675 2571 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6489153.
ACCESSION AR258675
VERSION AR258675.1 GI:27309061
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2571)
AUTHORS Wei, M.-H., Ketchum, K.A., Beasley, E.M. and DiFrancesco, V.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6489153-A 1 03-DEC-2002;
FEATURES
source
1..2571
/organism="unknown"
/mol_type="genomic DNA"

Query Match	100.0%	Score 2571;	DB 6;	Length 2571;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2571;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGGGCCGAGGGCGCGTGCCTGAGCGCCCATGGCCTTTCGCCCGCGCTCTTGGCGCG 60		
DB	1	CGGGCCGAGGGCGCGTGCCTGAGCGCCCATGGCCTTTCGCCCGCGCTCTTGGCGCG 60		
QY	61	GCACCTGTGGGGCGCGTGCCTGAGCGCGCGGGGTCTGCGTGGGGCCCATGGCTCCGCC 120		
DB	61	GCACCTGTGGGGCGCGTGCCTGAGCGCGCGGGGTCTGCGTGGGGCCCATGGCTCCGCC 120		
QY	121	GTGGCGCTTTCGCTTGGAGCTTCCGACTGCACCTTGGCTCTTTCGCCCTAGGGCGCGA 180		
DB	121	GTGGCGCTTTCGCTTGGAGCTTCCGACTGCACCTTGGCTCTTTCGCCCTAGGGCGCGA 180		
QY	181	CGCCCGCGCGACGACAGCGCCCGACCGCCCGCTGGCGCGCTGTTGGGCGCCCGCGA 240		
DB	181	CGCCCGCGCGACGACAGCGCCCGACCGCCCGCTGGCGCGCTGTTGGGCGCCCGCGA 240		
QY	241	GGCGACTACTCGCTGTGGTGGTCCCGTGCACCGCGCGCTGGGGGCGCGGCTCGG 300		

Db	241	GGCGAGCTACTGCGCTGTGCGTGCCCGGTGACCCCGGACGCGCGTGC	1321	CAGAGAAAAGGTCCTGCAGACAGTATTAAAGCCCTAATCCAGAATAGTTT	1380
Qy	301	GGCGGGGGCTGCACAGCGCTGTGTCACAGCTGCGCGCGGCCCTTCCAGCGGTG	1381	GTTACTCTGGCCAGGTGCCACGTCCTAACTAGATTAGATGTTTCTTTGAAACATCTACATCC	1440
Db	301	GGCGGGGGGGCTGCACAGCGCTGTGTCACAGCTGCGCGCGGCCCTTCCAGCGGTG	1381	GTTACTCTGGCCAGGTGCCACGTCCTAACTAGATTAGATGTTTCTTTGAAACATCTACATCC	1440
Qy	361	CCAGCTGCTCAGGCTGCTGTGTTACTGCGCGGGCGGCGAGGCGGGCGGCGACAGCAAGG	1441	ACCAATTGTTATGCAAGTGTCCCAAATTTCTGTTCTACAAGCATGTTGTGTGCAGAAAA	1500
Db	361	CCAGCTGCTCAGGCTGCTGTGTTACTGCGCGGGCGGCGAGGCGGGCGGCGACAGCAAGG	1441	ACCAATTGTTATGCAAGTGTCCCAAATTTCTGTTCTACAAGCATGTTGTGTGCAGAAAA	1500
Qy	421	CTTCTGCTGCGGACCCCTGTGATGACCTGTGACCCCGGACGCGCTGCTCGAGTGGT	1501	CTGGAGACCAGGCACTTTAAATTTTACTTTCAGCCATCGTACCCCTCTTCTGACTGATGGAC	1560
Db	421	CTTCTGCTGCGGACCCCTGTGATGACCTGTGACCCCGGACGCGCTGCTCGAGTGGT	1501	CTGGAGACCAGGCACTTTAAATTTTACTTTCAGCCATCGTACCCCTCTTCTGACTGATGGAC	1560
Qy	481	GGGCGCTCCGAGGAGCACCGCCGCACTTGGGCGGAGTTCCAGGCGCCACCCGCGCGG	1561	CGTCACTACAAAGTCCCTCTCATCATGTTCCAGTGAGAGGCGCAGGATGCTTCTTCTTC	1620
Db	481	GGGCGCTCCGAGGAGCACCGCCGCACTTGGGCGGAGTTCCAGGCGCCACCCGCGCGG	1561	CGTCACTACAAAGTCCCTCTCATCATGTTCCAGTGAGAGGCGCAGGATGCTTCTTCTTC	1620
Qy	541	CCAGCTGTGCGACGCGCTCTGGAGGTGCAAGACGCGAGCGCGCTGCGAGTGGGCTGGCG	1621	TGGCATAGTAAACATTTTCTTGGAACATATGTTTCACTTAACTACCTACCAATATCTGGA	1680
Db	541	CCAGCTGTGCGACGCGCTCTGGAGGTGCAAGACGCGAGCGCGCTGCGAGTGGGCTGGCG	1621	TGGCATAGTAAACATTTTCTTGGAACATATGTTTCACTTAACTACCTACCAATATCTGGA	1680
Qy	601	ACAGTGTGCTCCCGTCCCGGAGCCCGCTGACCCGCTGTCGCGAGCTTGCACAGTTC	1681	AGACCTGTCTTACTCAGACAGCAACCAGGTGTACAGAAGCAGCAGACAGATCTTCAGAT	1740
Db	601	ACAGTGTGCTCCCGTCCCGGAGCCCGCTGACCCGCTGTCGCGAGCTTGCACAGTTC	1681	AGACCTGTCTTACTCAGACAGCAACCAGGTGTACAGAAGCAGCAGACAGATCTTCAGAT	1740
Qy	661	CGTGTCTTCCCGGACCGGGAAGCCCGCGGCGGCTTTTGGAGGAGTGTACCTCTTTAT	1741	CAGCAGGGAGACCCCGGAGGCTTCTCTTCTTCTACACTGCGATGCGATGATGAGATCGTGAC	1800
Db	661	CGTGTCTTCCCGGACCGGGAAGCCCGCGGCGGCTTTTGGAGGAGTGTACCTCTTTAT	1741	CAGCAGGGAGACCCCGGAGGCTTCTCTTCTTCTTCTACACTGCGATGCGATGATGAGATCGTGAC	1800
Qy	721	TCCTGAAGCCCGGCGAGTGTTCACCTGTGTCACCTGTGTCACCTGTGTCACCTGTGTC	1801	ATGCCACACATTTGCTTCCACATCTGGTTGTCACCTGCTCATGATGGCTCGCTGCATCT	1860
Db	721	TCCTGAAGCCCGGCGAGTGTTCACCTGTGTCACCTGTGTCACCTGTGTCACCTGTGTC	1801	ATGCCACACATTTGCTTCCACATCTGGTTGTCACCTGCTCATGATGGCTCGCTGCATCT	1860
Qy	781	AAAGTTCAGGTTGTTGCCATCGAAGACTGGATGCGACCGGTAAACACACGCTGACCCA	1861	CCCTCAGTCCCAAATTTCTAGTAGCAAGTGTCTCTGTCAGAGGCTGTCTATGTGTCTGGC	1920
Db	781	AAAGTTCAGGTTGTTGCCATCGAAGACTGGATGCGACCGGTAAACACACGCTGACCCA	1861	CCCTCAGTCCCAAATTTCTAGTAGCAAGTGTCTCTGTCAGAGGCTGTCTATGTGTCTGGC	1920
Qy	841	GTCAAGTGGAGATTCACCTAAGGCTGCTCTTAAAGTCAACACCTCTTGCATTTGCCA	1921	TGCCAAGGAGCACTCTCGAGAGGCAATTTTGGGTAAAGAACACTTACAAAGAGGCAAT	1980
Db	841	GTCAAGTGGAGATTCACCTAAGGCTGCTCTTAAAGTCAACACCTCTTGCATTTGCCA	1921	TGCCAAGGAGCACTCTCGAGAGGCAATTTTGGGTAAAGAACACTTACAAAGAGGCAAT	1980
Qy	901	GTGAGGAAGATCTTTGATGATGAACCAACTATCATTTAGAGAGCTTTTACTTTTGGG	1981	TGATCTTGTCTGTGAGGCTCAGAGGCTTTTGTATAGGCTTCTGATGCTATTTCAAAAGAC	2040
Db	901	GTGAGGAAGATCTTTGATGATGAACCAACTATCATTTAGAGAGCTTTTACTTTTGGG	1981	TGATCTTGTCTGTGAGGCTCAGAGGCTTTTGTATAGGCTTCTGATGCTATTTCAAAAGAC	2040
Qy	961	CAATTATTTTGGCTCCGAAATAGCTAAAGATCTGCCAAATCTCTGTGATTGTAGA	2041	ATTCAAGCCCAAGATGCTTCAACTGCAATATATACCAACTTCTCTGAAATTTATTTTCTT	2100
Db	961	CAATTATTTTGGCTCCGAAATAGCTAAAGATCTGCCAAATCTCTGTGATTGTAGA	2041	ATTCAAGCCCAAGATGCTTCAACTGCAATATATACCAACTTCTCTGAAATTTATTTTCTT	2100
Qy	1021	CAGTACTGCGACAGCACGCGCACTATGCAATAGCCACTGAGGTGAGTGGGGTCTCCA	2101	ATTATATTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2160
Db	1021	CAGTACTGCGACAGCACGCGCACTATGCAATAGCCACTGAGGTGAGTGGGGTCTCCA	2101	ATTATATTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2160
Qy	1081	GCACTGCCCGCCAGCCCATCCCTGTGTACCACTGCGGCGAGAGCACTGCTCAAACTGA	2161	TCTGAACTGGATGATATCAATTTAGCCAAATCCAGTAAATTTATTTATTTATTTATTTAT	2220
Db	1081	GCACTGCCCGCCAGCCCATCCCTGTGTACCACTGCGGCGAGAGCACTGCTCAAACTGA	2161	TCTGAACTGGATGATATCAATTTAGCCAAATCCAGTAAATTTATTTATTTATTTATTTAT	2220
Qy	1141	CCCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2221	AATATGTTTCTCAGCAGTAGAGCTATGATTTATTTAAAGTGGAGTGAACCAAGCTA	2280
Db	1141	CCCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2221	AATATGTTTCTCAGCAGTAGAGCTATGATTTATTTAAAGTGGAGTGAACCAAGCTA	2280
Qy	1201	CATGAGAGAACCGAGGAAGAGCAACTTTCAGGCGCAACAGTGTGTTTCTGCAAAAGT	2281	AATGCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	2340
Db	1201	CATGAGAGAACCGAGGAAGAGCAACTTTCAGGCGCAACAGTGTGTTTCTGCAAAAGT	2281	AATGCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	2340
Qy	1261	AGAAATGCTTACAGCGGATGAGAAATCTTGGCTGCGATGCTGCTGCTGCTGCTGCTGCT	2341	CAAGTGGATCTTGGAGTGGGATTTCTTGGTAAATTTATCTTGCATCTGAAATGCTCATGAT	2400
Db	1261	AGAAATGCTTACAGCGGATGAGAAATCTTGGCTGCGATGCTGCTGCTGCTGCTGCTGCT	2341	CAAGTGGATCTTGGAGTGGGATTTCTTGGTAAATTTATCTTGCATCTGAAATGCTCATGAT	2400
Qy	1321	CAGAGAAAAGTCTCTGACAGACAGTATTAAAGCCCTAATCCAGAATAGTTT	2401	TACATATGAATCGCTTTCAGATATCTTTAGACAGAAAAAGTAGCTGATGAGGGGAA	2460
Db	1321	CAGAGAAAAGTCTCTGACAGACAGTATTAAAGCCCTAATCCAGAATAGTTT	2401	TACATATGAATCGCTTTCAGATATCTTTAGACAGAAAAAGTAGCTGATGAGGGGAA	2460

Db 1027 TCGTACCCCTCTTCTGACTGATGACCCGTCATCAAAAGGTCCTCTCATCATGTTCCAG 1086
QY 1595 TGAGAGCCGAGGATTGCTTTCTTCTGCGCATAGTAAACATTTCTTGGAACATATGTTT 1654
Db 1087 TGAGAGCCGAGGATTGCTTTCTTCTGCGCATAGTAAACATTTCTTGGAACATATGTTT 1146
QY 1655 CACTTAATCACTACCAATATCTGGAAGACCTGTCTTACTCAGACGACACAGGTTGTACA 1714
Db 1147 CACTTAATCACTACCAATATCTGGAAGACCTGTCTTACTCAGACGACACAGGTTGTACA 1206
QY 1715 GAAGCAGCAGACAAGATCTTCCAGATFCAGAGGAGACCCCGAGGCTCTGTTCTCCTTA 1774
Db 1207 GAAGCAGCAGACAAGATCTTCCAGATFCAGAGGAGACCCCGAGGCTCTGTTCTCCTTA 1266
QY 1775 CACTGCGATGCTGATGAGATCGTGCATCGCCACATTTGGCTTTCTTCCACATCTGTTGCA 1834
Db 1267 CACTGCGATGCTGATGAGATCGTGCATCGCCACATTTGGCTTTCTTCCACATCTGTTGCA 1326
QY 1835 CTGCTCATGATGGCTCGCTGCATCTCCTCAGTCGCAAAATCTAGTAGCCAAAGTGTTC 1894
Db 1327 CTGCTCATGATGGCTCGCTGCATCTCCTCAGTCGCAAAATCTAGTAGCCAAAGTGTTC 1385
QY 1895 TGCAGAGGCTGCTATGCTGCTGCTGCCAAGGAGACACTCTCTGCAGAGCCATTTTGG 1954
Db 1386 TGCAGAGGCTGCTATGCTGCTGCTGCCAAGGAGACACTCTCTGCAGAGCCATTTTGG 1444
QY 1955 GTAAGGAACACATTACAAGAGGATTTGATCTGCTGAGGCTCAGAGCCCTTTTGTAT 2014
Db 1445 GTAAGGAACACATTACAAGAGGATTTGATCTGCTGAGGCTCAGAGCCCTTTTGTAT 1504
QY 2015 AGGCTCTCATGCTCATTAAGACATTCAGCCCAAGATGCTCCAACTGCAAAATATACC 2074
Db 1505 AGGCTCTCTGA-GTCAATATAAAGACATTCAGCCCAAGATGCTCCAACTGCAAAATATACC 1563
QY 2075 AACCTTCTCTGAATATATTTGCTTATTTATTTATTTCTTTCTTTTCTTAAGAAATG 2134
Db 1564 AA-CTTCTCTGAATATA-TTGTCTATTTATTTATTTCTTTCTTTTCTTAAGTA--TG 1619
QY 2135 GCTCTGAATAGATGACATTTTCCATCTGAATCTGAATCATATCATTTAGCCAAATCCAG 2194
Db 1620 GCTCTGAATAGATGACATTTTCCAT-TGAATCTGAATGCAATTTCAATTTAGCCAAATCCAG 1678
QY 2195 TAATTTATTTATTAATCTATACATATATGTTTCTTCTCAGCATAGGAGCTATGTTAT 2254
Db 1679 TAATTTATTTATTAATCTATACATATATGTTTCTTCTCAGCATAGGAGCTATGTTAT 1738
QY 2255 TAATTAAGAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTATTTTCTATTACACAA 2314
Db 1739 TAATTAAGAGTGGAGTCAAAA-GCTAAATGCAATGT--GTGTGATTTTCTATTACACAA 1794
QY 2315 ACTTAATTTGCTGTTTAATTAAGTTCAAGTGGATCTTGGATGGGATTTCTT-GGTAAA 2373
Db 1795 A--TTAATTTGCTGTTTAATTAAGTACA--GGGATCTTGGATGGGATTTCTTGGGTAAA 1850
QY 2374 TTATCTGCATTTGAATGCTCATGATACATATGAATCGCTTTGACATATCTTTAGAC 2433
Db 1851 TTATCTGCATTTGAATGCTCTGATATACATATGAATTCGCTTGGATATCTTTTGGAC 1909
QY 2434 AGAAAAAGTAGCTAGTGGAGGGGAAATATATAGAGCTTGTGTGCTTTAGGAGTAGCT 2493
Db 1910 AGAAAAAGTAGTGGAGGGGAAATATATAGAGCTTGTGTGCTTTAGGAGTAGCT 1969

RESULT 4
LOCUS CQ490198
DEFINITION Sequence 22065 from Patent WO0160860.
ACCESSION CQ490198
VERSION CQ490198.1 GI:41455817
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schlegel, R., Endege, W.O. and Monahan, J.E.
Genes differentially expressed in human prostate cancer and their use
Patent: WO 0160860-A 22065 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
Location/Qualifiers
1..2199
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 61.4%; Score 1579.6; DB 6; Length 2199;
Best Local Similarity 97.1%; Pred. No. 9.1e-247;
Matches 1748; Conservative 0; Mismatches 34; Indels 18; Gaps 13;
QY 695 GTTTGGAGGAGTACCTCTTATTTCTGAAGCCCGGGCAGTGTGACCTGTGCGAC 754
Db 187 GCTTGTGCACAGTGTACCTCTTATTTCTGAAGCCCGGGCAGTGTGACCTGTGCGAC 246
QY 755 CAGTCCCAAAACAGATCCAGAAAGAAAGTTTCCAGGTTTGGTCCCATCGAAGGACTGGAT 814
Db 247 CAGTCCCAAAACAGATCCAGAAAGAAAGTTTCCAGGTTTGGTCCCATCGAAGGACTGGAT 306
QY 815 GCCACGGGTAAACCAACCGGTGACCCAGTCAGTGGCAGATTCACTTAAGGCTGTCTCTTA 874
Db 307 GCCACGGGTAAACCAACCGGTGACCCAGTCAGTGGCAGATTCACTTAAGGCTGTCTCTTA 366
QY 875 AAGTCACACCCCTCTTGTATTTGGCCAGTGGAGGAAGATCTTTGATGATGAACCAACTATC 934
Db 367 AAGTCACACCCCTCTTGTATTTGGCCAGTGGAGGAAGATCTTTGATGATGAACCAACTATC 426
QY 935 ATTAGAAGAGCTTTTACTCTTTGGCAATTATATTGGCCCTCCGAATAGCTAAAGAA 994
Db 427 ATTAGAAGAGCTTTTACTCTTTGGCAATTATATTGGCCCTCCGAATAGCTAAAGAA 486
QY 995 TCTGCCAAATCTCTGTGATTGTAGACAGGTACTGGCAGCAGCCGCCACCTATGCCATA 1054
Db 487 TCTGCCAAATCTCTGTGATTGTAGACAGGTACTGGCAGCAGCCGCCACCTATGCCATA 546
QY 1055 GCCACTGAGGTGAGTGGGGTCTCCAGACCTGCCGCCAGCCCATCACCTGTGTACCG 1114
Db 547 GCCACTGAGGTGAGTGGGGTCTCCAGACCTGCCGCCAGCCCATCACCTGTGTACCG 606
QY 1115 TGGCCAGAGGACCTCTCAAACTGACCTTATCTCTGTGCTCACTGTGAGTCTCTGAGGAG 1174
Db 607 TGGCCAGAGGACCTCTCAAACTGACCTTATCTCTGTGCTCACTGTGAGTCTCTGAGGAG 666
QY 1175 AGTTTGCAGAGCTCGAGGGCGGGCATGGAGAGACCCAGGAAAGAGCAAGCAACTTGAG 1234
Db 667 AGTTTGCAGAGCTCGAGGGCGGGCATGGAGAGACCCAGGAAAGAGCAAGCAACTTGAG 726
QY 1235 GCCACAGTGTGTTTCTGTCAAAAGTAGAATGTCCTTACCAGCGATGGAGATCTGTGC 1294
Db 727 GCCACAGTGTGTTTCTGTCAAAAGTAGAATGTCCTTACCAGCGATGGAGATCTGTGC 786
QY 1295 TGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGGTCCTGCAGACGATTAAGCCCTA 1354
Db 787 TGCCATGTGTTGATGCCAGCCCTCCAGAGAAAGGTCCTGCAGACGATTAAGCCCTA 846
QY 1355 ATCCAGAAATAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGTCTAACTAGATT 1414
Db 847 ATCCAGAAATAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGTCTAACTAGATT 906
QY 1415 AGATGTTGTTGAACATCTACATCCCATTTGTTATGAGTGTTCCTCAATTTCTGTT 1474
Db 907 AGATGTTGTTGAACATCTACATCCCATTTGTTATGAGTGTTCCTCAATTTCTGTT 966
QY 1475 CTACAAGCATGTTGTGGCAGAAACTGGAGACCGACCTCTTAATTTTACTTCAGCCA 1534

[illegible]

RESULT 5					
CQ494933					
LOCUS	CQ494933	2199 bp	DNA	Linear	PAT 30-JAN-2004
DEFINITION	Sequence 26800 from Patent WO0160860.				
ACCESSION	CQ494933				
VERSION	CQ494933.1	GI:41460552			

[illegible]

Db 730 CGTGGTCTTCGAGACGGGAAGCGCCGCGGCGTCTTTGGAGGAGTGTAACCTCCTTTAT 789
Qy 721 TCCTGAAGCCCGGAGAGTGTGACCTGGTGCACAGTGCCCAAAACAGATCCAGAAAGG 780
Db 790 TCCTGAAGCCCGGAGAGTGTGACCTGGTGCACAGTGCCCAAAACAGATCCAGAAAGG 849
Qy 781 AAAGTTCAGAGTGTGTGACATGAAGGACTGGATGCCAGGTTAAACACAGGTGACCCA 840
Db 850 AAAGTTCAGAGTGTGTGACATGAAGGACTGGATGCCAGGTTAAACACAGGTGACCCA 909
Qy 841 GTACGTGGCAGATCACTTAAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTGGCCA 900
Db 910 GTACGTGGCAGATCACTTAAGGCTGTCTCTTAAAGTCAACACCTCTTGCATTGGCCA 969
Qy 901 GTGAGGAGATCTTTGATGATGAACCACTATCAATTAGAGAGCTTTTACTCTTTGGG 960
Db 970 GTGAGGAGATCTTTGATGATGAACCACTATCAATTAGAGAGCTTTTACTCTTTGGG 1029
Qy 961 CAATTATATTGTGGCTCCGAAATAGCTAAAGATCTGCCAAATCTCCTGTGATTGTAGA 1020
Db 1030 CAATTATATTGTGGCTCCGAAATAGCTAAAGATCTGCCAAATCTCCTGTGATTGTAGA 1089
Qy 1021 CAGTACTGGACAGCAGCCACCTATGACATAGCCACTGAGGTGAGTGGGGTCTCCA 1080
Db 1090 CAGTACTGGACAGCAGCCACCTATGACATAGCCACTGAGGTGAGTGGGGTCTCCA 1149
Qy 1081 GCACCTGCCCCAGCCCATCACCTGTGTACCATGTCAGAGGACCTGTCTAAACCTGA 1140
Db 1150 GCACCTGCCCCAGCCCATCACCTGTGTACCATGTCAGAGGACCTGTCTAAACCTGA 1209
Qy 1141 CCTTATCTCTGTCTCACTGTGAGTCTCTGAGGAGAGGTTGACAGGCTGCAGGCGCGGG 1200
Db 1210 CCTTATCTCTGTCTCACTGTGAGTCTCTGAGGAGAGGTTGACAGGCTGCAGGCGCGGG 1269
Qy 1201 CATGGAGAACACAGGAGAGAGCAACTTGAGGCGCAACAGTGTGTTTCTCAAAAGGT 1260
Db 1270 CATGGAGAACACAGGAGAGAGCAACTTGAGGCGCAACAGTGTGTTTCTCAAAAGGT 1329
Qy 1261 AGAAATGTCTTACAGCGGATGAGATCTCTGCTGCGCATGTGGTTGATGCGACCGCTC 1320
Db 1330 AGAAATGTCTTACAGCGGATGAGATCTCTGCTGCGCATGTGGTTGATGCGACCGCTC 1389
Qy 1321 CAGAGAAAAGGTCTCTGACAGACAGTATTAAAGCCTTAATCCAGATAGTTTATGTAACCGTA 1380
Db 1390 CAGAGAAAAGGTCTCTGACAGAGTATTAAAGCCTTAATCCAGATAGTTTATGTAACCGTA 1449
Qy 1381 GTTACTCTGGCCAGGTGCCACGTCTAACTAGATTAGATGTTTGAACATCTACATCC 1440
Db 1450 GTTACTCTGGCCAGGTGCCACGTCTAACTAGATTAGATGTTTGAACATCTACATCC 1509
Qy 1441 ACCATTGTATTGACAGTGTTCCTCAA 1466
Db 1510 ACCATTGTATTGACAGTGTTCCTCAA 1535

RESULT 8
AX814308
LOCUS AX814308 1343 bp DNA linear PAT 05-DEC-2003
DEFINITION Sequence 1 from Patent WO03064642.
ACCESSION AX814308
VERSION AX814308.1 GI:39103548
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Liou, J.R.
REFERENCE
AUTHORS Regulation of human thymidylate kinase
TITLE Patent: WO 03064642-A 1 07-AUG-2003;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..1343
source

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.0%; Score 1336.6; DB 6; Length 1343;
Best Local Similarity 99.7%; Pred. No. 2.8e-207;
Matches 1339; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 32 ATGSCCTTCGCGCCCGGCTCCTCGCGGGCCACTGTGCGGGCGCTGCTCGGGCGGGC 91
Db 1 ATGSCCTTCGCGCCCGGCTCCTCGCGGGCCACTGTGCGGGCGCTGCTCGGGCGGGC 60
Qy 92 GGGGTCTGCGCTGGGGCCATGGCTCCGCGCTTCGCTCTGAGCTTCCCGACTGC 151
Db 61 GGGGTCTGCGCTGGGGCCATGGCTCCGCGCTTCGCTCTGAGCTTCCCGACTGC 120
Qy 152 ACCCTGGCTCACTTCGCGCTAGGGCCGACGCCCGCGGACGACGAGCGCCCGACCCC 211
Db 121 ACCCTGGCTCACTTCGCGCTAGGGCCGACGCCCGCGGACGACGAGCGCCCGACCCC 180
Qy 212 CGCTGCGCGGCTGTTGGGGCGCCCGGACGCGAGCTACTCGCTGTCGCTGCCGTGACC 271
Db 181 CGCTGCGCGGCTGTTGGGGCGCCCGGACGCGAGCTACTCGCTGTCGCTGCCGTGACC 240
Qy 272 CCGGACGCGCGCTGCGGGCGCCCGGCTCCGGCGCGGCTGCACACGCGCTGCTGCAC 331
Db 241 CCGGACGCGCGCTGCGGGCGCCCGGCTCCGGCGCGGCTGCACACGCGCTGCTGCAC 300
Qy 332 CAGCTGCGCGCGCGCGCTTCAGCGGTGCGAGCTGCTGAGGCTGCTCTGCTAGTGGCG 391
Db 301 CAGCTGCGCGCGCGCGCTTCAGCGGTGCGAGCTGCTGAGGCTGCTCTGCTAGTGGCG 360
Qy 392 GGGCGCGCGCGCGCGCCAGCAAGGCTTCTGCTGCGCGCACCCCTCGATGACCT 451
Db 361 GGGCGCGCGCGCGCGCCAGCAAGGCTTCTGCTGCGCGCACCCCTCGATGACCT 420
Qy 452 GACACCGCGCAAGCGTGTCTGAGCTGCTGGGCGCTGCCAGGAGCACACACCGCGCAC 511
Db 421 GACACCGCGCAAGCGTGTCTGAGCTGCTGGGCGCTGCCAGGAGCACACACCGCGCAC 480
Qy 512 TTGGGGAGTTCGAGGCGGCGCGCGCGCGCGCGCTGCGAGCTGCTGGGAGGTGCAA 571
Db 481 TTGGGGAGTTCGAGGCGGCGCGCGCGCGCGCGCTGCGAGCTGCTGGGAGGTGCAA 540
Qy 572 GACGCGAGCGGCTGCGAGGTGGGCTGCGCACAGTCTGCGCGCTGCCGAGCGCCGCTG 631
Db 541 GACGCGAGCGGCTGCGAGGTGGGCTGCGCACAGTCTGCGCGCTGCCGAGCGCCGCTG 600
Qy 632 CACCGGTGCTGCGAGCTTGGCTTCTTTCGAGTCTTTCGAGTCTTTCGAGTCTTTCGAGT 691
Db 601 CACCGGTGCTGCGAGCTTGGCTTCTTTCGAGTCTTTCGAGTCTTTCGAGTCTTTCGAGT 660
Qy 692 GCGTTTTGAGGAGTGTACCTCTTATTCCTGAGCGCGGCGAGTCTTTCGAGTCTTTCGAGT 751
Db 661 GCGTTTTGAGGAGTGTACCTCTTATTCCTGAGCGCGGCGAGTCTTTCGAGTCTTTCGAGT 720
Qy 752 GACGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 811
Db 721 GACGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Qy 812 GATGCGCGGGTAAACCAACGCTGACCCAGTCCAGTGGCAGATTCATTTAAGCTGTCTTC 871
Db 781 GATGCGCGGGTAAACCAACGCTGACCCAGTCCAGTGGCAGATTCATTTAAGCTGTCTTC 840
Qy 872 TTAAAGTACACCCCTCTTTCGATTCGCGAGTGGAGAGATCTTTGATGATGAACCACT 931
Db 841 TTAAAGTACACCCCTCTTTCGATTCGCGAGTGGAGAGATCTTTGATGATGAACCACT 900
Qy 932 ATCATTAGAGAGCTTTTACTCTTTGGGCAATTATTTGTCCTTCCGAAATAGCTAA 991
Db 901 ATCATTAGAGAGCTTTTACTCTTTGGGCAATTATTTGTCCTTCCGAAATAGCTAA 960

[illegible]


```
/rpt_family="MIR"
11761..11851
/rpt_family="MERL_type"
12251..12472
/rpt_family="L1"
12956..13250
/rpt_family="GA-rich"
13251..13674
/rpt_family="MarL"
13675..13711
/rpt_family="GA-rich"
13869..14320
/rpt_family="L1"
14623..14682
/rpt_family="L1"
14683..15380
/rpt_family="L1"
16147..16223
/rpt_family="MIR"
16504..16530
/rpt_family="(CAGAGA)n"
17250..17408
/rpt_family="MIR"
17912..18102
/rpt_family="MIR"
18596..19161
/rpt_family="L1"
19837..20226
/rpt_family="ERV1"
19882..20401
/note="similar to EST AA781159 (NID:g2840490) aj24c01.s1"
21312..21474
/rpt_family="MIR"
21497..21604
/rpt_family="MIR"
21651..22045
/note="similar to EST AI220331 (NID:g3802534) qg90e09.x1"
21914..22187
/rpt_family="L1"
22338..22465
/rpt_family="GA-rich"
22498..22559
/rpt_family="MIR"
23060..23231
/rpt_family="L2"
23335..23631
/rpt_family="Alu"
23711..24048
/rpt_family="L1"
24050..24168
/rpt_family="L2"
24891..25221
/note="similar to EST D63066 (NID:g967991)"
25260..25415
/rpt_family="MIR"
25484..25555
/rpt_family="Mariner"
25558..25680
/rpt_family="L2"
25969..26139
/rpt_family="L2"
26445..26531
/rpt_family="MIR"
26600..26627
/rpt_family="AT-rich"
27490..27594
/rpt_family="MIR"
27983..28373
/note="similar to EST AW075990 (NID:g6030988) xa82f12.x1"
28112..28406
/rpt_family="Alu"
28421..28673
/note="similar to EST AI874110 (NID:g5548159) wm49b06.x1"

misc_feature 28538..28962
/note="similar to EST AI955710 (NID:g5748020) wt37e05.x1"
misc_feature 28541..28968
/note="similar to EST AI056212 (NID:g3330078) oy94h05.x1"
repeat_region 29077..29381
/rpt_family="Alu"
repeat_region 29700..29817
/rpt_family="GA-rich"
repeat_region 29819..29989
/rpt_family="L1"
repeat_region 30022..30479
/rpt_family="L1"

Query Match 44.4%; Score 1141.4; DB 9; Length 118808;
Best Local Similarity 98.2%; Pred. No. 1.9e-175;
Matches 1228; Conservative 0; Mismatches 16; Indels 7; Gaps 7;

QY 1243 TGTGTTTCGTCAAAAGGTAGAAATGCTTACACAGCGATGGAGAAATCTGGCTGCCATGT 1302
Db 49421 TTTGCCTGTCCCAGGGTAGAAATGCTTACCAGCGATGGAGAAATCTGGCTGCCATGT 49362

QY 1303 GGTGATGCCAGCCCTCCAGAGAAAGGTCTTCGACAGACAGTATTAAAGCCTAATCAGAA 1362
Db 49361 GGTGATGCCAGCCCTCCAGAGAAAGGTCTTCGACAGACAGTATTAAAGCCTAATCAGAA 49302

QY 1363 TAGTTTTAGTGAACCGTAGTTACTCTGGCAGGTGCCACGTCTAACTAGATTAGATTGTG 1422
Db 49301 TAGTTTTAGTGAACCGTAGTTACTCTGGCAGGTGCCACGTCTAACTAGATTAGATTGTG 49242

QY 1423 TTTGAAACATCTACATCCACCATTTGTTATGAGTGTTCCTCAAAATTTCTGTTCTACAGC 1482
Db 49241 TTTGAAACATCTACATCCACCATTTGTTATGAGTGTTCCTCAAAATTTCTGTTCTACAGC 49182

QY 1483 ATGTTGTGGCAGAGAACTGGAGACAGGCATCTTAATTTTACTTCAGCATCGTACCC 1542
Db 49181 ATGTTGTGGCAGAGAACTGGAGACAGGCATCTTAATTTTACTTCAGCATCGTACCC 49122

QY 1543 TCTTCTGACTGATGGAGCCCGTCATCAAAAGGTCCCTCTCATCATGTTCCAGTGAGAGGC 1602
Db 49121 TCTTCTGACTGATGGAGCCCGTCATCAAAAGGTCCCTCTCATCATGTTCCAGTGAGAGGC 49062

QY 1603 CAGCATGCTTTCTTCCTGGCATAGTAAACATTTTCTTGGAAACATATGTTCACTTAAT 1662
Db 49061 CAGCATGCTTTCTTCCTGGCATAGTAAACATTTTCTTGGAAACATATGTTCACTTAAT 49002

QY 1663 CACTACCAATATCTGGAAGACCTGCTTACTCAGACAGCACAGGTGTACAGAGCAGC 1722
Db 49001 CACTACCAATATCTGGAAGACCTGCTTACTCAGACAGCACAGGTGTACAGAGCAGC 48942

QY 1723 AGACAAGATCTTCCAGATCAGAGGGAGACCCCGAGCCCTCTGCTTCTCTTACACTGGCA 1782
Db 48941 AGACAAGATCTTCCAGATCAGAGGGAGACCCCGAGCCCTCTGCTTCTCTTACACTGGCA 48882

QY 1783 TGTGATGAGATGATGACATGCCCATATGCTTCTTCCACATCTGTTGACCTCGTCAT 1842
Db 48881 TGTGATGAGATGATGACATGCCCATATGCTTCTTCCACATCTGTTGACCTCGTCAT 48822

QY 1843 GATGGGTCTGCTGCATCTCCCTCAGTCCCAAAATTTAGTAGCCAAAGTGTTCCTCAGAGG 1902
Db 48821 GATGGGTCTGCTGCATCTCCCTCAGTCCCAAAATTTAGTAGCCAAAGTGTTCCTCAGAGG 48763

QY 1903 CTGTCATGTCCTGCTGCCAAGGGAACATCTCCTGAGAGCCATTTTGGGTAAGGAA 1962
Db 48762 CTGTCATGTCCTGCTGCCAAGGGAACATCTCCTGAGAGCCATTTTGGGTAAGGAA 48704

QY 1963 CACTTACAAAGAGGCATTGATCTTGCTGCTGAGGCTCAGAGCCCTTTTGTAGGCTTCT 2022
Db 48703 CACTTACAAAGAGGCATTGATCTTGCTGCTGAGGCTCAGAGCCCTTTTGTAGGCTTCT 48644

QY 2023 GATGTCATTCATAAAGACATTCAGAGCCAAAGATGCTCCAACCTGCAAAATATACCAACTTCT 2082
Db 48643 GA-GTCATATATAAAGACATTCAGAGCCAAAGATGCTCCAACCTGCAAAATATACCAACTTCT 48585
```


SOURCE	Unknown.	ORGANISM	Unclassified.
	Unknown.		Unclassified.
REFERENCE	1 (bases 1 to 20966)	AUTHORS	Wei, M.-H., Kechum, K.A., Beasley, E.M. and DiFrancesco, V.
	Isolated human kinase proteins		
JOURNAL	Patent: US 664087-A 3 16-DEC-2003;	TITLE	Location/Qualifiers
	1. .20966		/organism="unknown"
FEATURES	Source	ORIGIN	Query Match 44.0%; Score 1130.4; DB 6; Length 20966;
			Best Local Similarity 98.1%; Pred. No. 1.1e-173;
			Matches 1228; Conservative 0; Mismatches 16; Indels 8; Gaps 8;
QY	1243 TGTGTTTCGTCAAAGGTAGAAATGTCCTACACGCGGATGGAGAATCCTGGCTGCCATGT	Db	1302
Db	17724 TTTGCTGTCTCCAGGGTAGAAATGTCCTACACGCGGATGGAGAATCCTGGCTGCCATGT		17783
QY	1303 GGTGATGCCAGCCCTCCAGAGAAAGGTCTCTGCAGACAGTAGTATTAGCCCTAATCCAGAA	Db	1362
Db	17784 GGTGATGCCAGCCCTCCAGAGAAAGGTCTCTGCAGACAGTAGTATTAGCCCTAATCCAGAA		17843
QY	1363 TAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGTCTAACTAGATTAGATTGTG	Db	1422
Db	17844 TAGTTTGTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGTCTAACTAGATTAGATTGTG		17903
QY	1423 TTGGAACATCTACATCCACATTTGTATGTCAGTGTTCCTCAAAATTTCTGTTCTACAAAGC	Db	1482
Db	17904 TTTGAAACATCTACATCCACATTTGTATGTCAGTGTTCCTCAAAATTTCTGTTCTACAAAGC		17963
QY	1483 ATGTGTGTGTGCAGAAACTGGACACGAGCATCTTAATTTTACTTCAGCCATCGTACCC	Db	1542
Db	17964 ATGTGTGTGTGCAGAAACTGGACACGAGCATCTTAATTTTACTTCAGCCATCGTACCC		18023
QY	1543 TC-TTCTGACGTAGTGAACCGTCTATCAAAAGTTCCTCTCATCATGTTCCAGTGAGAGG	Db	1601
Db	18024 TCTTTCTGACGTAGTGAACCGTCTATCAAAAGTTCCTCTCATCATGTTCCAGTGAGAGG		18083
QY	1602 CCAGCGATGCTTTCTCTGGCATAGTAACATTTCTTGGAACATATGTTTCACTTAA	Db	1661
Db	18084 CCAGCGATGCTTTCTCTGGCATAGTAACATTTCTTGGAACATATGTTTCACTTAA		18143
QY	1662 TCACCTACAAATATCTGGAACCTGTCTTACTCAGACAGACACAGGTGTACAGAACGAG	Db	1721
Db	18144 TCACCTACAAATATCTGGAACCTGTCTTACTCAGACAGACACAGGTGTACAGAACGAG		18203
QY	1722 CAGACAGATCTTCAGATCAGCGGAGACCCCGAGCCCTCTGCTTCTCTACACTGGC	Db	1781
Db	18204 CAGACAGATCTTCAGATCAGCGGAGACCCCGAGCCCTCTGCTTCTCTACACTGGC		18263
QY	1782 ATGCTGATGAGATCGTGACATGCCACATTTGCTTCTTCCACATCTGGTTGCACTCGCA	Db	1841
Db	18264 ATGCTGATGAGATCGTGACATGCCACATTTGCTTCTTCCACATCTGGTTGCACTCGCA		18323
QY	1842 TGATGGGCTCGTGCATCTCCCTCAGTCCCAAAATCTAGTACCCAAAGTGTCTCTCCAGAG	Db	1901
Db	18324 TGATGGGCTCGTGCATCTCCCTCAGTCCCAAAATCTAGTACCCAAAGTGTCTCTCCAGAG		18382
QY	1902 GCTGTCTATGTCTGCTGCTGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAAGGA	Db	1961
Db	18383 GCTGTCTATGTCTGCTGCTGCCCAAGGACACTCTCTGAGAGCCATTTTGGGTAAAGGA		18441
QY	1962 ACACCTTACAAAGAGCATTTGATCTGTGTCGAGGCTCAGAGCCCTTTTGTATAGCTTTC	Db	2021
Db	18442 ACACCTTACAAAGAGCATTTGATCTGTGTCGAGGCTCAGAGCCCTTTTGTATAGCTTTC		18501
QY	2022 TGATGTCTATCAATAAGACATTTCAAGCCCAAGATGCTTCCAACTGTGCAAAATATACCACTTTC	Db	2081
Db	18502 TGA-GTCAATATATAAGACATTTCAAGCCCAAGATGCTTCCAACTGTGCAAAATATACCACTTTC		18560

QY	2082 TCTGAATTATATTTTCTGTTTATTTATTTATTTCTTTCTTTTCTTAAAGATTTGGCTCTGA	2141
Db	18561 TCTGAATTATATTTTCTGTTTATTTATTTATTTCTTTCTTTTCTTAAAGTA-TGGCTCTGA	18619
QY	2142 ATAGAATGCACATTTTCCATCTGAATCTGAATGCATATCATTTAGCAATCCAGTAATTTA	2201
Db	18620 ATAGAATGCACATTTTCCAT-TGAATCTGAATGCATTTTATTTAGCAATCCAGTAATTTA	18678
QY	2202 TTTATATTAATCTATACATAAATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTA	2261
Db	18679 TTTATATTAATCTATACATAAATATGTTTCTCAGCATAGGAGCTATGATTCATTAATTA	18738
QY	2262 AAGTGGAGTCAAAACCGCTAAATGCAATGTTTGTGTATTTTTCATTACACAACTAAT	2321
Db	18739 AAGTGGAGTCAAAACCGCTAAATGCAATGTTTGTGTATTTTTCATTACACAACTAAT	18798
QY	2322 TTGCTCTGTTTAAATAGTTCAGTGGATCTTGGAGTGGGATTTCTTGGTAAATATCTTG	2381
Db	18799 TTGCTCTGTTTAAATAGTAC-AGTGGATCTTGGAGTGGGATTTCTTGGTAAATATCTTG	18857
QY	2382 CACTTGAATGTCATCATGATTACATATGAAATCGCTTTGACATATCTTTAGACAGAAAAA	2441
Db	18858 CACTTGAATGTCATCATGATTACATATGAAATCGCTTTGACATATCTTTAGACAGAAAAA	18917
QY	2442 GTAGCTGAGTGGGGGAAATTTATAGAGCTTGTGACCTTTAGGGAGTAGCT	2493
Db	18918 GTAGCTGAGTGGGGGAAATTTATAGAGC-TGTGTGACTTTAGGGAGTAGGT	18968
RESULT 13		
LOCUS	CO733692	1215 bp DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 19626 from Patent WO02068579.	
ACCESSION	CO733692	
VERSION	CO733692.1	GI:42318470
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens	
REFERENCE	1	
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof	
JOURNAL	Patent: WO 02068579-A 19626 06-SEP-2002;	
FEATURES	PE Corporation (NY) (US)	
source	Location/Qualifiers	
	1. .1215	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Query Match	41.4%; Score 1065.2; DB 6; Length 1215;	
Best Local Similarity	89.8%; Pred. No. 3.5e-163;	
Matches 1212; Conservative	0; Mismatches 3; Indels 135; Gaps 1;	
QY	32 ATGGCTTCGCGCGCGGCTCTCGCGGGGCACTGTCTGGGCGCGCTCTCGGGCGGCG	91
Db	1 ATGGCTTCGCGCGCGGCTCTCTCGCGGGGCACTGTCTGGGCGCGCTCTCGGGCGGCG	60
QY	92 GGGGTCTCGGTGGGGCCATGGCTCCGCGCTGCCCTTCGTCCTCGAGCTTCCCACTGC	151
Db	61 GGGGTCTCGGTGGGGCCATGGCTCCGCGCGCGCTTCGTCCTCGAGCTTCCCACTGC	120
QY	152 ACCCTTGGTCACTTCCCTTAGCGCCGACGCCCCCGGCGAGCGAGCCCCGACCCC	211
Db	121 ACCCTTGGTCACTTCCCTTAGCGCCGACGCCCCCGGCGAGCGAGCCCCGACCCC	180
QY	212 CCGCTTGGCGGCTGTCTGGGCGCGCGGAGCGAGCTACTCGCTGTGGCTGCCGTGACC	271
Db	181 CCGCTTGGCGGCTGTCTGGGCGCGCGGAGCGAGCTACTCGCTGTGGCTGCCCT	234

```
QY 272 CCGGACGCCGCTCGGGGCGCGGTCGGGCGCGGCTGACACGAGCGCTCTGCTGAC 331
Db 235 -----
QY 332 CAGCTGCGCGCGCGCGCTTCCAGCGGTGTCAGCTGCTCAGGCTGCTCTGCTACTGCGCG 391
Db 235 -----
QY 392 GCGGCGCAGGCGCGCGCGACACAAAGGCTTCTGCTGCGCGACCCCTCTGGATGACCT 451
Db 235 -----GCGGCGCGCGCGCGCTTCTGCTGCGCGACCCCTCTGGATGACCT 285
QY 452 GACACCGCGCAAGCGCTGCTCGAGCTGCTGGCGCCCTGCGAGGAGCACACGCCCGAC 511
Db 286 GACACCGCGCAAGCGCTGCTCGAGCTGCTGGCGCCCTGCGAGGAGCACACGCCCGAC 345
QY 512 TTGGCGGAGTTCGAGGCGCGACCCGCGCGCGAGCTGTGGCAGCGCTCTGGAGGTGCAA 571
Db 346 TTGGCGGAGTTCGAGGCGCGACCCGCGCGCGAGCTGTGGCAGCGCTCTGGAGGTGCAA 405
QY 572 GACGCGAGGCGCTGCAAGGTGGGTGCGCACAGTGTGCGCGGACCCCGCGCTG 631
Db 406 GACGCGAGGCGGTGCAAGGTGGGTGCGCACAGTGTGCGCGGACCCCGCGCTG 465
QY 632 CACCGGTGTGCGAGACTTGGCCAGTTCGCTGCTTCCGCGACCGGAGCGCGCGG 691
Db 466 CACCGGTGTGCGAGACTTGGCCAGTTCGCTGCTTCCGCGACCGGAGCGCGCGG 525
QY 692 GCGGTTTGGAGGAGTGTACTCTCTTATTCCTGAAGCCGCGGAGTGTGACCTGGTC 751
Db 526 GCGGTTTGGAGGAGTGTACTCTCTTATTCCTGAAGCCGCGGAGTGTGACCTGGTC 585
QY 752 GACAGTGCCTCAAAACAGATCCAGAAAGAAAGTTCAGGTTGTGCGATCGAAGGACTG 811
Db 586 GACAGTGCCTCAAAACAGATCCAGAAAGAAAGTTCAGGTTGTGCGATCGAAGGACTG 645
QY 812 GATGCCACGGGTAAACACCGGTGACCGAGTTCAGTGCAGATTCACCTAAGCTGTCTC 871
Db 646 GATGCCACGGGTAAACACCGGTGACCGAGTTCAGTGCAGATTCACCTAAGCTGTCTC 705
QY 872 TTAAGTCAACACCTCTTGCATTGGCCAGTGGAGGAAGATCTTGTATGATGAACCAACT 931
Db 706 TTAAGTCAACACCTCTTGCATTGGCCAGTGGAGGAAGATCTTGTATGATGAACCAACT 765
QY 932 ATCAITAGAGACTTTTACTCTTTGGGCAATTATTTGGCTCTCCGAAATAGCTAA 991
Db 766 ATCAITAGAGAGCTTTTACTCTTTGGGCAATTATTTGGCTCTCCGAAATAGCTAA 825
QY 992 GAATCTGCCAAATCTCTGTGATTGTAGACAGTACTGGCACACGCGGACCTATGCC 1051
Db 826 GAATCTGCCAAATCTCTGTGATTGTAGACAGTACTGGCACACGCGGACCTATGCC 885
QY 1052 ATAGCCACTGAGTGTGGGTCTCCAGCAGCTGCGCCCGCCACCTACCTGTGTAC 1111
Db 886 ATAGCCACTGAGTGTGGGTCTCCAGCAGCTGCGCCCGCCACCTACCTGTGTAC 945
QY 1112 CAGTGCCACAGAGACTGTCAAACCTGACCTTATCTGTGTGTCTATGTAGTCTGTAG 1171
Db 946 CAGTGCCACAGAGACTGTCAAACCTGACCTTATCTGTGTGTCTATGTAGTCTGTAG 1005
QY 1172 GAGAGTTGACAGGCTGACAGGCGCGGCGATGCGAGAGCAGGAGGAGGAGGACTT 1231
Db 1006 GAGAGTTGACAGGCTGACAGGCGCGGCGATGCGAGAGCAGGAGGAGGAGGACTT 1065
QY 1232 GAGGCCAACAGTGTGTTTCTCAAAAGGTAGAAATGTCTTACCGCGGATGGAATCCT 1291
Db 1066 GAGGCCAACAGTGTGTTTCTCAAAAGGTAGAAATGTCTTACCGCGGATGGAATCCT 1125
QY 1292 GGTGCTCATGTGTTGATGCGACCCCTCCAGAGAAAGGTCTGTCAGACAGTATTAAAG 1351
Db 1126 GGTGCTCATGTGTTGATGCGACCCCTCCAGAGAAAGGTCTGTCAGACAGTATTAAAG 1185
```

```
QY 1352 CTATCCAGAATAGTTTTTACTGAACCGTAG 1381
Db 1186 CTATCCAGAATAGTTTTTACTGAACCGTAG 1215

RESULT 14
AK127983
LOCUS
DEFINITION
  Homo sapiens cDNA FLJ46098 fis, clone TESTI2021654, weakly similar
  to Mus musculus thymidylate kinase family LPS-inducible member
  (Tyk1).
ACCESSION
  AK127983
KEYWORDS
  oligo capping; fis (full insert sequence).
SOURCE
  Homo sapiens
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
  Tanai,H., Matanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,
  Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
  Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
  Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
  Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
  Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
  and Isogai,T.
  NEDO human cDNA sequencing project
  Unpublished
  2 (bases 1 to 1609)
  Isogai,T. and Yamamoto,J.
  Direct Submission
  Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
  Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
  (E-mail:genomics@ri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
  NEDO human cDNA sequencing project supported by Ministry of
  Economy, Trade and Industry of Japan; cDNA full insert sequencing:
  Research Association for Biotechnology (RAB); cDNA library
  construction: Helix Research Institute (HRI) (supported by Japan
  Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
  HRI, and Biotechnology Center, National Institute of Technology and
  Evaluation; clone selection for full insert sequencing: HRI and
  RAB; annotation: Reverse Proteomics Research Institute, HRI and
  RAB.

FEATURES
    source
    1..1609
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="TESTI2021654"
        /tissue_type="testis"
        /clone_lib="TESTI2"
        /note="cloning vector: pME18SFL3"
    212..1123
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="BAC87217.1"
        /gb_xref="GI:34535132"
        /translation="MAFARLLRGLSPGLGRGVGAGAMAPRRFVLELPDCTLAH
        FALGDADGADADPRLAALLGPPERSYSLCVPTPDGCGARVRAARLHORLLHQL
        RRGPRQCLRLCYCPGGAGGACQGLLDLPDLPDTRQALLLELGCAGCEAPRPH
        LGEFADPRGLQWLWEVDQRQLQVCAQVVPPEPLFVPLPSVYFPPEA
        ARAVLEECTSFIPPEARVLDLVDCPQKIQKQFVAIEGLDATGKTIVTQSVADSL
        KAVPQITLCTSGQRTCSNLTLSCSL"

ORIGIN
  Query Match 34.6%; Score 889.2; DB 9; Length 1609;
  Best Local Similarity 83.4%; Pred. No. 1.5e-134;
  Matches 1121; Conservative 0; Mismatches 3; Indels 220; Gaps 1;

  QY 1 CGGGGCGGAGGCGGCTGCTGAGCGGCCCATGCTCGCCCGCGCTCTGCGCGG 60
  Db 181 CGGGGCGGAGGCGGCGTCTGAGCGGCCCATGCTCTGCGCGCGCTCTGCGCGG 240
```

Qy	61	GCACATGTCGGGCGCGTGTCTGGGCGGCGCGGGGTCTGGGCTGGGGCCATGGCTCCGCC	120
Db	241	GCACATGTCGGGCGCGTGTCTGGGCGGCGCGGGGTCTGGGCTGGGGCCATGGCTCCGCC	300
Qy	121	GTGCGGCTTCTGTGAGCTTCCGACTGCACCTTGGCTCACTTCGCCCTTAGGCGCCGA	180
Db	301	GCGCGGCTTCTGTGAGCTTCCGACTGCACCTTGGCTCACTTCGCCCTTAGGCGCCGA	360
Qy	181	CGCCCCGGGCGACGACGCCCCCGACCCCGCCCTGTGGGCGGCTGTGGGGCCCCCGGA	240
Db	361	CGCCCCGGGCGACGACGCCCCCGACCCCGCCCTGTGGGCGGCTGTGGGGCCCCCGGA	420
Qy	241	GCGCAGCTACTGCTGTGCTGTCGCTGACCGCGGACCGCGGCTGCGGGGCCCGGCTCGG	300
Db	421	GCGCAGCTACTGCTGTGCTGTCGCTGACCGCGGACCGCGGCTGCGGGGCCCGGCTCGG	480
Qy	301	GCGCGCGGCTGACACGAGCGCTGCTGTCACACGCTGCGCGCGGCGCCCTTCACGCGGT	360
Db	481	GCGCGCGGCTGACACGAGCGCTGCTGTCACACGCTGCGCGCGGCGCCCTTCACGCGGT	540
Qy	361	CCAGCTGTCTAGGCTGTCTGCTACTGCTCGGGGCGGCGAGCGCGGCGGCGCACAGG	420
Db	541	CCAGCTGTCTAGGCTGTCTGCTACTGCTCGGGGCGGCGAGCGCGGCGGCGCACAGG	600
Qy	421	CTTCTGCTGCGGACCCCTTGATGACCTGACACCGGCAAGCGCTGCTCGAGCTGCT	480
Db	601	CTTCTGCTGCGGACCCCTTGATGACCTGACACCGGCAAGCGCTGCTCGAGCTGCT	660
Qy	481	GCGCGCTGCCAGGAGCACACGCGGCTGCTGGGCGAGTTCAGAGCGGCGGCGCGG	540
Db	661	GCGCGCTGCCAGGAGCACACGCGGCTGCTGGGCGAGTTCAGAGCGGCGGCGGCGG	720
Qy	541	CCAGCTGTGGAGCGCTCTGAGGTCGAGAGCGGAGGCGGCTGAGGTGGGCTCGCG	600
Db	721	CCAGCTGTGGAGCGCTCTGAGGTCGAGAGCGGAGGCGGCTGAGGTGGGCTCGCG	780
Qy	601	ACAGTGTGTCGCGCTGCCGAGCGGCGCTGACCCGCGTGTGCGACAGTTCGCCAGTTC	660
Db	781	ACAGTGTGTCGCGCTGCCGAGCGGCGCTGACCCGCGTGTGCGACAGTTCGCCAGTTC	840
Qy	661	CTGTGTCTCCGAGCGGAGCGCGCGGCGGCTTTGAGGAGTGTACTCTCTTAT	720
Db	841	CTGTGTCTTTCCGAGCGGAGCGCGCGGCGGCTTTGAGGAGTGTACTCTCTTAT	900
Qy	721	TCTGAAGCGCGGCGAGTGTGACTGCTGACACGAGTGCACCAACAGATCCGAAAGG	780
Db	901	TCTGAAGCGCGGCGAGTGTGACTGCTGACACGAGTGCACCAACAGATCCGAAAGG	960
Qy	781	AAAGTTCAGGTTTGCCATCGAAGAGTGTGATCCGAGGTAACACAGGTAACCA	840
Db	961	AAAGTTCAGGTTTGCCATCGAAGAGTGTGATCCGAGGTAACACAGGTAACCA	1020
Qy	841	GTGAGGCGAGATTACATTAAAGCTGTCTCTTAAAGTCAACACCTCTTGATGGCA	900
Db	1021	GTGAGGCGAGATTACATTAAAGCTGTCTCTTAAAGTCAACACCTCTTGATGGCT	1047
Qy	901	GTGGAGGAAGATCTTTGATGATGAACCAACTATCATTAGAGAGCTTTTACTCTTGGG	960
Db	1048	-----	1047
Qy	961	CAATTATTTGGGCTCCGAATAGCTAAAGATCTGCAAAATCTCTGTGATTGTAGA	1020
Db	1048	-----	1047
Qy	1021	CAGGTACTGACACAGCAGGCGCCATTCATGACCTAGCTAGGTGAGTGGGGTCTCA	1080
Db	1048	-----	1047
Qy	1081	GCACCTGCCCCAGCCCATCACCTGTGTACAGTGGCGCAGAGGACCTCTCAACCTGA	1140
Db	1048	-----	1100

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 125 Row: b Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 15805017.

FEATURES

source

1. .3028
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:66848 IMAGE:6401867"
 /tissue_type="Mouse, brain 12.5 dpc"
 /clone_lib="NIH BMAP_FCO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1. .3028
 /gene="Tyki"
 /note="synonym: TDKI"
 /db_xref="locusID:22169"
 /db_xref="MGI:99830"
 41. .1294
 /gene="Tyki"
 /codon_start=1
 /product="Tyki protein"
 /protein_id="AAH57565.1"
 /db_xref="GI:34785825"
 /db_xref="LocusID:22169"
 /db_xref="MGI:99830"
 /translation="MARPRFTVELPDCSLTHFVLGDAIDHRDARLAELLGPGRSYA
 LCVPLAPEGCGPRVQARVHRLLOLRGPIRLCOLSLKLYGPGDOAGEAQHGL
 LRQCDHPDRRLDLQLLGSQCOEAPRLAEFOADSOGLLWRLWELQSDROVQVDC
 CVLPAQPHLPHLLPDLLNSAVFQDRDAARVLEECRSFIPPEARVAVLDVQCPREVQ
 KGRFQVIAI EGLDARGTKITLITVSSELSKAVLIQSPPCISQWRKIFDDEPTIIRAF
 YSLGNIVASEI AKESNFVYVDRYWHSTATVAITEVSGGLQYLPAPHPVYQWPG
 DLLKPDVLLLTIVNSERVRLRGQGEKTEEALEANNVFRKVENTYQRMENPSC
 HLVDASPSRETVLQKVLELIQSSGR"

gene

CDS

ORIGIN

Query Match 34.0%; Score 873; DB 10; Length 3028;
 Best Local Similarity 77.2%; Pred. No. 6.5e-132;
 Matches 1105; Conservative 0; Mismatches 310; Indels 16; Gaps 3;
 70 GGGGCCCTGCTGGGGCGGCGGGGCTGCGCTGGGGCCATGCTCGCGCTCCGCTT 129
 Db 1 GCGGAGCTGTGGGGCGACACAGGCTGTCATGGGACCATGCTCGGGCGGCGCTT 60
 130 CGTCTGAGCTTCCGACTGCACTCGGCTCACTTCGCCCTAGGCGCGACGCCCGCG 189
 Db 61 CACTGTGGAGTGCAGATGCTCCCTGACTCACTTCGTCTCTGG-----GGA 108
 190 CGAGGAGCGCCCGACCCCGCTGGCGGCGCTGTGGGGCCCGCGAGCGAGCTA 249
 Db 109 TGCAACAGACACAGGATGACGCTTGGCAGAGCTGCTGGGCCCGCGAGGCGAGCTA 168
 250 CTGCTGTGGTGGTACCCCGAGCGCGGCTGGGGCGGCGCGGCTCGGGCGGCGCG 309
 Db 169 CGGCTGTGTGGCCCTGGCTCAGCGGAAGGCTGGGGCCCGGGTGGAGGGCGCCG 228
 310 GTGACACAGCGCTGTGACACAGCTGCGCGCGCGCCCTTCAGCGGTGCCAGCTGCT 369
 Db 229 GGTGACCATCGCTGTGCGAGAGCTGCGCGGGTCCCTTACAGAGATGCCAAGTGA 288
 370 CAGCTGCTGTGCTACTGCGGGCGGCGAGCGCGGGCGGCGGCGAGAGCTTCTGCT 429
 Db 289 CAAGCTTCTGGGCTATGCTCGGGCGATCAGCGCGCGGAAGCCAGCATGCTCTCTGCT 348
 430 GCGGACCCCTGGATGACCTGACACCGGCAAGCGCTGCTCGAGCTGTGGGCGCTG 489
 Db 349 GCGGACCCCTGGACACCCCGGACACTCGGCGGCACTTGTCTCAGCTTCTGGGCTCTG 408
 490 CCAGGAGGACACCGCCCGCACTTGGGCGAGTTCCAGGCGGACCGCGCGCGCAGCTGTG 549
 Db 409 CCAGGAGGCGGCGCGCGCGAGTTGGCGGAGTTCCAGGCGGACTCCAGGGTTGCTGTG 468

Search completed: October 22, 2004, 18:08:31
 Job time : 10852 secs

This Page Blank (uspto)